SINGULAR SOLUTIONS IN OPTIMAL CONTROL: SECOND ORDER CONDITIONS AND A SHOOTING ALGORITHM[1,2]

M. SOLEDAD ARONNA

Abstract. In this article we study optimal control problems for systems that are affine in one part of the control variable. Finitely many equality and inequality constraints on the initial and final state are considered. We investigate singular solutions for this class of problems. First, we obtain second order necessary and sufficient conditions for weak optimality. Afterwards, we propose a shooting algorithm and show that the sufficient condition above-mentioned is also sufficient for the local quadratic convergence of the algorithm.

1. Introduction

The purpose of this paper is to investigate optimal control problems governed by ordinary differential equations that are affine in one part of the control variable. This class of system includes both the totally affine and the nonlinear cases.

Many models that correspond to this framework can be found in practice and, in particular, in the existing literature. Among these we can mention: the Goddard's problem [25] in 3 dimensions analyzed in Martinon et al. [11], other models concerning the motion of a rocket in Lawden [38], Bell and Jacobson [9], Goh [27, 31], Oberle [46], Azimov [8] and Hull [33]; an hydrothermal electricity production problem studied in Bortolossi et al. [13] and Aronna et al. [6], the problem of atmospheric flight considered by Oberle in [48], and an optimal production process in Cho et al. [17] and Maurer at al. [41].

The subject of second order optimality conditions for these partially affine problems have been studied by Goh in [28, 29, 27, 31], Dmitruk in [21], Dmitruk and Shishov in [22], Bernstein and Zeidan [10], and Maurer and Osmolovskii [42]. The first works were by Goh, who introduced a change of variables in [28] and used it to obtain optimality conditions in [28, 26, 27], always assuming uniqueness of the multiplier. The necessary conditions we present imply those by Goh [26] when there is only one multiplier. Recently,

Key words and phrases. optimal control, singular control, second order optimality condition, weak optimality, shooting algorithm, Gauss-Newton method.

*This article was published as the INRIA Research Report Nr. 7764.

†This work is supported by the European Union under the 7th Framework Programme FP7-PEOPLE-2010-ITN Grant agreement number 264735-SADCO.

Dmitruk and Shishov [22] analysed the quadratic functional associated with the second variation of the Lagrangian function and provided a set of necessary conditions for the nonnegativity of this quadratic functional. Their results are consequence of a second order necessary condition we present. In [21] Dmitruk proposed, without proof, necessary and sufficient conditions for a problem having a particular structure: the affine control variable applies to a term depending only on the state variable, i.e. the affine and nonlinear controls are 'uncoupled'. This hypothesis is not used in our work. The conditions established here coincide with those suggested in Dmitruk [21] when the latter are applicable. In [10], Bernstein and Zeidan derived a Riccati equation for the singular linear-quadratic regulator, which is a modification of the classical linear-quadratic regulator where only some components of the control enter quadratically in the cost function. All of these four articles use Goh's Transformation to derive their conditions; we use this transformation as well. On the other hand, in [42] Maurer and Osmolovskii gave a sufficient condition for a class of problems having one affine control subject to bounds and such that it is bang-bang at the optimal solution. This structure is not studied here since no control constraints are considered, i.e. our optimal control is suppose to be totally singular.

Regarding second order optimality conditions, we provide a pair of necessary and sufficient conditions for weak optimality of totally singular solutions. These conditions are 'no gap' in the sense that the sufficient condition is obtained from the necessary one by strengthening an inequality. We do not assume uniqueness of multiplier.

Among the applications of the shooting method to the numerical solution of partially affine problems we can mention the articles Oberle [45, 48] and Oberle-Taubert [49]. In these articles the authors use a generalization of the algorithm that Maurer [39] suggested for totally affine systems. These works present interesting implementations of a shooting-like method to solve partially affine control problems having bang-singular or bang-bang solutions and, in some cases, running-state constraints are considered. No result on convergence is given in these articles.

In this paper we propose a shooting algorithm which can be also used to solve problems with bound on the controls. Our algorithm is an extension of the method for totally affine problems in Aronna et al. [7]. We give a theoretical support to this method, by showing that the second order sufficient condition above-mentioned ensures the local quadratic convergence of the algorithm.

The article is organised as follows. In Section 2 we present the problem, the basic definitions and first order conditions. In Section 3 we give the tools for second order analysis and establish a second order necessary condition. We introduce Goh's Transformation in Section 4. In Section 5 we show a new necessary condition, and in Section 6 we give a sufficient one. A shooting algorithm is proposed in Section 7, and in Section 8 we prove

that the sufficient condition above-mentioned guarantees the local quadratic convergence of the algorithm.

Notations. We denote by h_t the value of function h at time t if h is a function that depends only on t, and by $h_{i,t}$ the ith component of h evaluated at t. Partial derivatives of a function h of (t,x) are referred as $D_t h$ and $D_x h$. When dealing with derivatives of higher order we may use the notation of type h_{xx} since it is not ambiguous. By $\mathbb{R}^k$ we denote the $k-$dimensional real space, i.e. the space of column real vectors of dimension k; and by $\mathbb{R}^{k,*}$ its corresponding dual space, which consists of $k-$dimensional row real vectors. By $L^p(0,T;\mathbb{R}^k)$ we mean the Lebesgue space with domain equal to the interval $[0,T] \subset \mathbb{R}$ and with values in $\mathbb{R}^k$. The notation $W^{q,s}(0,T;\mathbb{R}^k)$ refers to the Sobolev spaces (see Adams [1] for further details on Sobolev spaces).

2. Statement of the problem and assumptions

2.1. **Statement of the problem.** We study the optimal control problem (P) given by

$$J := \varphi_0(x_0, x_T) \to \min, \tag{1}$$

$$\dot{x}_t = \sum_{i=0}^{m} v_{i,t} f_i(x_t, u_t), \quad \text{a.e. on } [0,T], \tag{2}$$

$$\eta_j(x_0, x_T) = 0, \quad \text{for } j = 1\ldots, d_\eta, \tag{3}$$

$$\varphi_i(x_0, x_T) \le 0, \quad \text{for } i = 1, \ldots, d_\varphi. \tag{4}$$

Here $f_i : \mathbb{R}^{n+l} \to \mathbb{R}^n$ for $i = 0, \ldots, m$, $\varphi_i : \mathbb{R}^{2n} \to \mathbb{R}$ for $i = 0, \ldots, d_\varphi$, $\eta_j : \mathbb{R}^{2n} \to \mathbb{R}$ for $j = 1, \ldots, d_\eta$ and we put, in sake of simplicity of notation, $v_0 \equiv 1$ which is not a variable. The *nonlinear control* u belongs to $\mathcal{U} := L^\infty(0,T;\mathbb{R}^l)$, while by $\mathcal{V} := L^\infty(0,T;\mathbb{R}^m)$ we denote the space of *affine controls* v, and $\mathcal{X} := W^{1,\infty}(0,T;\mathbb{R}^n)$ refers to the state space. When needed, we write $w = (x,u,v)$ for a point in $\mathcal{W} := \mathcal{X} \times \mathcal{U} \times \mathcal{V}$. The hypothesis below is considered along all the article.

Assumption 2.1. All data functions have Lipschitz-continuous second derivatives.

A *trajectory* is an element $w \in \mathcal{W}$ that satisfies the state equation (2). If in addition, constraints (3) and (4) hold, we say that w is a *feasible trajectory* of problem (P).

Definition 2.2. A feasible trajectory $\hat{w} = (\hat{x}, \hat{u}, \hat{v}) \in \mathcal{W}$ is a *weak minimum* of (P) if there exists $\varepsilon > 0$ such that the cost function attains at $\hat{w}$ its minimum in the set of feasible trajectories $w = (x,u,v)$ satisfying

$$\|x - \hat{x}\|_\infty < \varepsilon, \quad \|u - \hat{u}\|_\infty < \varepsilon, \quad \|v - \hat{v}\|_\infty < \varepsilon.$$

In the sequel, we study a nominal feasible trajectory $\hat{w} = (\hat{x}, \hat{u}, \hat{v}) \in \mathcal{W}$. An element $\delta w \in \mathcal{W}$ is termed *feasible variation for* $\hat{w}$ if $\hat{w} + \delta w$ is feasible for (P). Take $\lambda = (\alpha, \beta, p)$ in $\mathbb{R}^{d_\varphi+1,*} \times \mathbb{R}^{d_\eta,*} \times W^{1,\infty}(0,T;\mathbb{R}^{n,*})$. Define the *pre-Hamiltonian* function

$$H[\lambda](x,u,v,t) := p_t \sum_{i=0}^{m} v_i f_i(x,u),$$

the *terminal Lagrangian* function $\ell : \mathbb{R}^{2n} \to \mathbb{R}$ by

$$\ell[\lambda](q) := \sum_{i=0}^{d_\varphi} \alpha_i \varphi_i(q) + \sum_{j=1}^{d_\eta} \beta_j \eta_j(q),$$

and the *Lagrangian* function

$$\text{Ł}[\lambda](w) := \ell[\lambda](x_0, x_T) + \int_0^T p_t \left(\sum_{i=0}^{m} v_{i,t} f_i(x_t, u_t) - \dot{x}_t \right) \mathrm{d}t. \tag{5}$$

We assume, in sake of simplicity, that whenever some argument of f_i, H, ℓ, Ł or their derivatives is omitted, they are evaluated at $\hat{w}$. Without loss of generality we suppose that

$$\varphi_i(\hat{x}_0, \hat{x}_T) = 0, \text{ for all } i = 1, \dots, d_\varphi. \tag{6}$$

2.2. **Lagrange multipliers.** We introduce here the concept of multiplier. The second order conditions that we prove in this article are expressed in terms of the second variation of the Lagrangian in (5) and the set of Lagrange multipliers associated with $\hat{w}$ that we define below.

Definition 2.3. An element $\lambda = (\alpha, \beta, p) \in \mathbb{R}^{d_\varphi+1,*} \times \mathbb{R}^{d_\eta,*} \times W^{1,\infty}(0,T;\mathbb{R}^{n,*})$ is a *Lagrange multiplier* associated with $\hat{w}$ if it satisfies the following conditions,

$$|\alpha| + |\beta| = 1, \tag{7}$$

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{d_\varphi}) \geq 0, \tag{8}$$

the function p is solution of the *costate equation*

$$-\dot{p}_t = D_x H[\lambda](\hat{x}_t, \hat{u}_t, \hat{v}_t, t), \tag{9}$$

and it satisfies the *transversality conditions*

$$\begin{aligned} p_0 &= -D_{x_0} \ell[\lambda](\hat{x}_0, \hat{x}_T), \\ p_T &= D_{x_T} \ell[\lambda](\hat{x}_0, \hat{x}_T), \end{aligned} \tag{10}$$

and the *stationarity* conditions

$$\begin{cases} D_u H[\lambda](\hat{x}(t), \hat{u}(t), \hat{v}(t), t) = 0, \\ D_v H[\lambda](\hat{x}(t), \hat{u}(t), \hat{v}(t), t) = 0, \end{cases} \quad \text{a.e. on } [0,T], \tag{11}$$

hold true. Denote by Λ the *set of Lagrange multipliers* associated with $\hat{w}$.

Recall the following well-known result.

Theorem 2.4. *If $\hat{w}$ is a weak minimum, the set Λ is non empty and compact.*

Proof. Regarding the existence of a Lagrange multiplier the reader is referred to [3, 37], [44, Thm. 2.1]. In order to prove the compactness, observe that p may be expressed as a linear continuous mapping of (α, β). Thus, since the normalization (7) holds, Λ is a finite-dimensional compact set. ☐

In view of previous result, note that Λ can be identified with a compact subset of $\mathbb{R}^s$, where $s := d_\varphi + d_\eta + 1$.

Given $(\bar{x}_0, \bar{u}, \bar{v}) \in \mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$, consider the *linearized state equation*

$$\dot{\bar{x}}_t = A_t\bar{x}_t + E_t\bar{u}_t + B_t\bar{v}_t, \quad \text{a.e. on } [0,T], \tag{12}$$

$$\bar{x}(0) = \bar{x}_0, \tag{13}$$

where

$$A_t := \sum_{i=0}^{m} \hat{v}_i D_x f_i(\hat{x}, \hat{u}), \quad E_t := \sum_{i=0}^{m} \hat{v}_i D_u f_i(\hat{x}, \hat{u}), \tag{14}$$

and $B : [0,T] \to \mathcal{M}_{n\times m}(\mathbb{R})$ such that for every $v \in \mathbb{R}^m$,

$$B_t v := \sum_{i=1}^{m} v_i f_i(\hat{x}_t, \hat{u}_t). \tag{15}$$

Here $\mathcal{M}_{n\times m}(\mathbb{R})$ refers to the space of $n \times m-$real matrices. Hence, the ith. column of B is $f_i(\hat{x}, \hat{u})$. The solution $\bar{x}$ of (12)-(13) is called *linearized state variable.*

2.3. **Critical cones.** We define now the sets of critical directions associated with $\hat{w}$, both in the $L^\infty-$ and the L^2-norm. Even if we are working with control variables in L^∞ and hence the control perturbations are naturally taken in L^∞, the second order analysis involves quadratic mappings and it is useful to extend them continuously to L^2.

Set $\mathcal{X}_2 := W^{1,2}(0,T;\mathbb{R}^n)$, $\mathcal{U}_2 := L^2(0,T;\mathbb{R}^l)$ and $\mathcal{V}_2 := L^2(0,T;\mathbb{R}^m)$. Put $\mathcal{W}_2 := \mathcal{X}_2 \times \mathcal{U}_2 \times \mathcal{V}_2$ for the corresponding product space. Given $\bar{w} \in \mathcal{W}_2$ satisfying (12)-(13), consider the *linearization of the endpoint constraints and cost function,*

$$D\eta_j(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T) = 0, \quad \text{for } j = 1, \ldots, d_\eta, \tag{16}$$

$$D\varphi_i(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T) \leq 0, \quad \text{for } i = 0, \ldots, d_\varphi. \tag{17}$$

Define the *critical cones* in $\mathcal{W}$ and $\mathcal{W}_2$ by

$$\mathcal{C} := \{\bar{w} \in \mathcal{W} : (12)\text{-}(13), (16)\text{-}(17) \text{ hold}\}, \tag{18}$$

$$\mathcal{C}_2 := \{\bar{w} \in \mathcal{W}_2 : (12)\text{-}(13), (16)\text{-}(17) \text{ hold}\}. \tag{19}$$

Lemma 2.5. *The critical cone $\mathcal{C}$ is a dense subset of $\mathcal{C}_2$.*

In order to prove previous lemma, recall the following technical result (see e.g. Dmitruk [20, Lemma 1] for a proof).

Lemma 2.6 (on density of cones)**.** *Consider a locally convex topological space X, a finite-faced cone $Z \subset X$, and a linear manifold Y dense in X. Then the cone $Z \cap Y$ is dense in Z.*

Proof. [of Lemma 2.5]
Set $X := \{\bar{w} \in \mathcal{W}_2 : (12)\text{-}(13)\,\text{hold}\}$, $Y := \{\bar{w} \in \mathcal{W} : (12)\text{-}(13)\,\text{hold}\}$, and $Z := \mathcal{C}_2$ and apply Lemma 2.6. □

3. Second order analysis

We begin this section by giving an expression of the second derivative of the Lagrangian function Ł, in terms of the derivatives of ℓ and H. We denote it by Ω. All the second order conditions we present are established in terms of either Ω or some transformed form of Ω. The main result of the current section is the necessary condition in Theorem 3.9, which is applied in Section 5 to get Theorem 5.3.

3.1. **Second variation.** Let us consider the quadratic mapping

(20)
$$\begin{aligned}\Omega[\lambda](\bar{x},\bar{u},\bar{v}) :=\ & \tfrac{1}{2}D^2\ell[\lambda](\hat{x}_0,\hat{x}_T)(\bar{x}_0,\bar{x}_T)^2 \\ &+ \int_0^T [\tfrac{1}{2}\bar{x}^\top Q[\lambda]\bar{x} + \bar{u}^\top F[\lambda]\bar{x} + \bar{v}^\top C[\lambda]\bar{x} + \tfrac{1}{2}\bar{u}^\top R_0[\lambda]\bar{u} + \bar{v}^\top K[\lambda]\bar{u}]\mathrm{d}t,\end{aligned}$$

where the involved matrices are, omitting arguments,

$$Q := H_{xx}, \quad F := H_{ux}, \quad C := H_{vx}, \quad R_0 := H_{uu}, \quad K := H_{vu}. \tag{21}$$

Recall the following notation: given two functions $h : \mathbb{R}^N \to \mathbb{R}^M$ and $k : \mathbb{R}^N \to \mathbb{R}^L$, we say the h is a *big-O* of k around 0 and denote it by

$$h(x) = \mathcal{O}(k(x)),$$

if there exists positive constants δ and M such that $|h(x)| \leq M|k(x)|$ for $|x| < \delta$. It is a *small-o* if M goes to 0 as $|x|$ goes to 0. Denote this by

$$h(x) = o(k(x)).$$

Lemma 3.1 (Lagrangian expansion)**.** *Let $w = (x,u,v) \in \mathcal{W}$ be a solution of (2), and set $\delta w = (\delta x, \delta u, \delta v) := w - \hat{w}$. Then for every multiplier $\lambda \in \Lambda$,*

$$Ł[\lambda](w) = Ł[\lambda](\hat{w}) + \Omega[\lambda](\delta x,\delta u,\delta v) + \tau[\lambda](\delta x,\delta u,\delta v) + \mathcal{R}(\delta x,\delta u,\delta v), \tag{22}$$

where the time variable is omitted in the sake of simplicity, τ is a cubic mapping given by

$$\tau[\lambda](\delta x,\delta u,\delta v) := \int_0^T [H_{vxx}[\lambda](\delta x,\delta x,\delta v) + 2H_{vux}[\lambda](\delta x,\delta u,\delta v) + H_{vuu}[\lambda](\delta u,\delta u,\delta v)]\,\mathrm{d}t,$$

and $\mathcal{R}$ satisfies the estimate

$$\mathcal{R}(\delta x,\delta u,\delta v) = \mathcal{O}(|(\delta x_0,\delta x_T)|^3) + (1 + \|v\|_1)\|(\delta x,\delta u)\|_\infty \mathcal{O}(\|(\delta x,\delta u)\|_2^2).$$

Proof. Omit the dependence on λ for the sake of simplicity. In order to achieve the expression (22) consider the second order Taylor representations below, written in a compact form,

$$\ell(x_0,x_T)=\ell+D\ell(\delta x_0,\delta x_T)+\tfrac{1}{2}D^2\ell(\delta x_0,\delta x_T)^2+\mathcal{O}(|(\delta x_0,\delta x_T)|^3), \tag{23}$$
$$f_i(x_t,u_t)=f_{i,t}+Df_i(\delta x_t,\delta u_t)+\tfrac{1}{2}D^2f_i(\delta x_t,\delta u_t)^2+\mathcal{O}(\|(\delta x,\delta u)\|_3^3), \tag{24}$$

where, whenever the argument is missing, the corresponding function is evaluated on the reference trajectory $\hat{w}$. Observe that the transversality conditions (10) and the costate equation (9) yield

$$D\ell\,(\delta x_0,\delta x_T)=-p_0\,\delta x_0+p_T\,\delta x_T=\int_0^T p\left[-\sum_{i=0}^m \hat{v}_i D_x f_i\,\delta x+\dot{\delta x}\right]\mathrm{d}t. \tag{25}$$

Recall the expression of the Lagrangian given in (5). Replacing $\ell(x_0,x_T)$ and $f_i(x,u)$ in (5) by their Taylor expansions (23)-(24) and using the identity (25) we get

$$\begin{aligned}
Ł(w)=Ł(\hat{w})&+\int_0^T[H_u\delta u+H_v\delta v]\mathrm{d}t+\Omega(\delta x,\delta u,\delta v)\\
&+\int_0^T[H_{vxx}(\delta x,\delta x,\delta v)+2H_{vux}(\delta x,\delta u,\delta v)+H_{vuu}(\delta u,\delta u,\delta v)]\,\mathrm{d}t\\
&+\mathcal{O}(|(\delta x_0,\delta x_T)|^3)+\|(\delta x,\delta u)\|_\infty\int_0^T p\sum_{i=0}^m v_i\,\mathcal{O}(\|(\delta x,\delta u)\|_2^2)\mathrm{d}t.
\end{aligned}$$

Finally, to obtain (22) use stationarity condition (11) and the compactness of Λ. $\square$

Remark 3.2. The last lemma yields the equality

$$\Omega[\lambda](\bar{w})=\tfrac{1}{2}D^2Ł[\lambda](\hat{w})\,\bar{w}^2. \tag{26}$$

3.2. **Second order necessary condition.** Recall the second order condition below.

Theorem 3.3 (Classical second order necessary condition)**.** *If $\hat{w}$ is a weak minimum of problem (P), then*

$$\max_{\lambda\in\Lambda}\Omega[\lambda](\bar{x},\bar{u},\bar{v})\geq 0,\quad \text{on } \mathcal{C}. \tag{27}$$

A proof of Theorem 3.3 can be found in Osmolovskii [50]. Nevertheless, for the sake of completeness, we give a proof here.

We shall write problem (P) in an abstract form and, therefore, we consider the functions

$$\bar{\eta}_j:\mathbb{R}^n\times\mathcal{U}\times\mathcal{V}\to\mathbb{R},\quad \bar{\eta}_j(x_0,u,v):=\eta_j(x_0,x_T), \tag{28}$$
$$\bar{\varphi}_i:\mathbb{R}^n\times\mathcal{U}\times\mathcal{V}\to\mathbb{R},\quad \bar{\varphi}_i(x_0,u,v):=\varphi_i(x_0,x_T), \tag{29}$$

where $x \in \mathcal{W}$ is the solution of (2) associated with (x_0, u, v). Hence, (P) can be written as the following problem in the space $\mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$,

$$\text{(AP)} \qquad \begin{aligned} &\min \bar{\varphi}_0(x_0, u, v);\\ &\text{s.t. } \bar{\eta}_j(x_0, u, v) = 0, \text{ for } j = 1, \dots, d_\eta,\\ &\qquad \bar{\varphi}_i(x_0, u, v) \leq 0, \text{ for } j = 1, \dots, d_\varphi. \end{aligned}$$

Notice that if $\hat{w}$ is a weak solution of (P) then $(\hat{x}_0, \hat{u}, \hat{v})$ is a local solution of (AP).

Definition 3.4. We say that the endpoint equality constraints are *qualified* if

$$D\bar{\eta}(\hat{x}_0, \hat{u}, \hat{v}) \text{ is onto from } \mathbb{R}^n \times \mathcal{U} \times \mathcal{V} \text{ to } \mathbb{R}^{d_\eta}. \tag{30}$$

When (30) does not hold, the constraints are *not qualified.*

The proof of Theorem 3.3 is divided in two cases: qualified and not qualified endpoint equality constraints. In the latter case the condition (27) follows easily and it is shown in Lemma 3.5 below. The proof for the qualified case is done by means of an auxiliary problem written in an abstract form and its dual.

Lemma 3.5. *If the equality constraints are not qualified then* (27) *holds.*

Proof. Observe that since $D\bar{\eta}(\hat{x}_0, \hat{u}, \hat{v})$ is not onto there exists $\beta \in \mathbb{R}^{d_\eta,*}$ with $|\beta| = 1$ such that $\sum_{j=1}^{d_\eta} \beta_j D\bar{\eta}_j(\hat{x}_0, \hat{u}, \hat{v}) = 0$ and consequently,

$$\sum_{j=1}^{d_\eta} \beta_j D\eta_j(\hat{x}_0, \hat{x}_T) = 0.$$

Set $\lambda := (p, \alpha, \beta)$ with $p \equiv 0$ and $\alpha = 0$. Then both λ and $-\lambda$ are in Λ. Observe that

$$\Omega[\lambda](\bar{x}, \bar{u}, \bar{v}) = \tfrac{1}{2} \sum_{j=1}^{d_\eta} \beta_j D^2\eta_j(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T)^2.$$

Thus, either $\Omega[\lambda](\bar{x}, \bar{u}, \bar{v})$ or $\Omega[-\lambda](\bar{x}, \bar{u}, \bar{v})$ is nonnegative. The desired result follows. $\square$

Let us now deal with the qualified case. Take a critical direction $\bar{w} = (\bar{x}, \bar{u}, \bar{v})$ and consider the problem in the variables $\zeta \in \mathbb{R}$ and $r = (r_{x_0}, r_u, r_v) \in \mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$ given by

$$(\mathrm{QP}_{\bar{w}}) \qquad \begin{aligned} &\min \zeta\\ &\text{s.t. } D\bar{\eta}(\hat{x}_0, \hat{u}, \hat{v})r + D^2\bar{\eta}(\hat{x}_0, \hat{u}, \hat{v})(\bar{x}_0, \bar{u}, \bar{v})^2 = 0,\\ &\qquad D\bar{\varphi}_i(\hat{x}_0, \hat{u}, \hat{v})r + D^2\bar{\varphi}_i(\hat{x}_0, \hat{u}, \hat{v})(\bar{x}_0, \bar{u}, \bar{v})^2 \leq \zeta, \quad i = 0, \dots, d_\varphi. \end{aligned}$$

Proposition 3.6. *Assume that $\hat{w}$ is a weak solution of (P) such that the endpoint equality constraints are qualified at $\hat{w}$. Let $\bar{w} \in \mathcal{C}$ be a critical direction. Then the problem* $(\mathrm{QP}_{\bar{w}})$ *is feasible and has nonnegative value.*

Proof. Step I. Let us first show feasibility. Since $D\bar{\eta}(\hat{x}_0,\hat{u},\hat{v})$ is onto, there exists $r \in \mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$ such that the equality constraint in $(\mathrm{QP}_{\bar{w}})$ is satisfied. Set

$$\zeta := \max_{0\leq i\leq d_\varphi} \{D\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})r + D^2\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})(\bar{x}_0,\bar{u},\bar{v})^2\}.$$

Then (ζ, r) is feasible for $(\mathrm{QP}_{\bar{w}})$.

Step II. Let us now prove that $(\mathrm{QP}_{\bar{w}})$ has nonnegative value. Suppose on the contrary that there is $(\zeta, r) \in \mathbb{R}\times\mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$ feasible for $(\mathrm{QP}_{\bar{w}})$ with $\zeta < 0$. We look now for a family of feasible solutions of (P) that we will denote by $\{r(\sigma)\}_\sigma$. It shall be defined for small positive σ and satisfy

$$r(\sigma) \to (\hat{x}_0,\hat{u},\hat{v}), \quad \bar{\varphi}_0(r(\sigma)) < \bar{\varphi}_0(\hat{x}_0,\hat{u},\hat{v}). \tag{31}$$

The existence of $\{r(\sigma)\}_\sigma$ will contradict the local optimality of $(\hat{x}_0,\hat{u},\hat{v})$. Consider hence

$$\tilde{r}(\sigma) := (\hat{x}_0,\hat{u},\hat{v}) + \sigma(\bar{x}_0,\bar{u},\bar{v}) + \tfrac{1}{2}\sigma^2 r.$$

Let $0 \leq i \leq d_\varphi$ and observe that

$$\begin{aligned}\bar{\varphi}_i(\tilde{r}(\sigma)) &= \bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v}) + \sigma D\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})(\bar{x}_0,\bar{u},\bar{v})\\ &\quad + \tfrac{1}{2}\sigma^2\left[D\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})r + D^2\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})(\bar{x}_0,\bar{u},\bar{v})^2\right] + o(\sigma^2)\\ &\leq \bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v}) + \tfrac{1}{2}\sigma^2\zeta + o(\sigma^2).\end{aligned} \tag{32}$$

Analogously,

$$\bar{\eta}(\tilde{r}(\sigma)) = o(\sigma^2). \tag{33}$$

Since $D\bar{\eta}(\hat{x}_0,\hat{u},\hat{v})$ is onto, there exists $r(\sigma)$ such that $\|r(\sigma)-\tilde{r}(\sigma)\|_\infty = o(\sigma^2)$ and $\bar{\eta}(r(\sigma)) = 0$. This follows by applying the Implicit Function Theorem to the mapping

$$(r,\sigma) \mapsto \bar{\eta}\left((\hat{x}_0,\hat{u},\hat{v}) + \sigma(\bar{x}_0,\bar{u},\bar{v}) + \tfrac{1}{2}\sigma^2 r\right) = \bar{\eta}(\tilde{r}(\sigma)).$$

On the other hand, by taking σ sufficiently small in estimate (32), we can obtain

$$\bar{\varphi}_i(r(\sigma)) < \bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v}), \tag{34}$$

since $\zeta < 0$. Hence $r(\sigma)$ is feasible for (AP) and verifies (31). This contradicts the optimality of $(\hat{x}_0,\hat{u},\hat{v})$. We conclude then that all the feasible solutions of $(\mathrm{QP}_{\bar{w}})$ have $\zeta \geq 0$, and therefore its value is nonnegative. □

We shall now go back to the proof of Theorem 3.3.

Proof. [of Theorem 3.3] The not qualified case is covered by Lemma 3.5 above. Hence, for this proof, assume that (30) holds.

Given $\bar{w} \in \mathcal{C}$, Proposition 3.6 implies that there cannot exist $(\zeta, r) \in \mathbb{R}\times\mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$ such that

$$\begin{cases} D\bar{\eta}(\hat{x}_0,\hat{u},\hat{v})r + D^2\bar{\eta}(\hat{x}_0,\hat{u},\hat{v})(\bar{x}_0,\bar{u},\bar{v})^2 = 0,\\ D\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})r + D^2\bar{\varphi}_i(\hat{x}_0,\hat{u},\hat{v})(\bar{x}_0,\bar{u},\bar{v})^2 \leq \zeta, \quad i = 0,\dots,d_\varphi,\\ \zeta < 0.\end{cases}$$

Therefore, the Dubovitskii-Milyutin Theorem (see [23]) guarantees the existence of $(\alpha,\beta)\in\mathbb{R}^{1+d_\varphi+d_\eta}$ satisfying,

$$\begin{gathered}\sum_{i=0}^{d_\varphi}\alpha_i D\bar\varphi_i(\hat x_0,\hat u,\hat v)+\sum_{j=1}^{d_\eta}\beta_j D\bar\eta_j(\hat x_0,\hat u,\hat v)=0,\\ \sum_{i=0}^{d_\varphi}\alpha_i D^2\bar\varphi_i(\hat x_0,\hat u,\hat v)(\bar x_0,\bar u,\bar v)^2+\sum_{j=1}^{d_\eta}\beta_j D^2\bar\eta_j(\hat x_0,\hat u,\hat v)(\bar x_0,\bar u,\bar v)^2\geq 0\end{gathered}\tag{35}$$

Let p be the solution of (9)-(10) associated with (α,β). Let us show that $\lambda:=(\alpha,\beta,p)$ is a Lagrange multiplier for $\hat w$. In fact, observe that

$$\begin{aligned}0&=\sum_{i=0}^{d_\varphi}\alpha_i D\bar\varphi_i(\hat x_0,\hat u,\hat v)(\bar x_0,\bar u,\bar v)+\sum_{j=1}^{d_\eta}\beta_j D\bar\eta_j(\hat x_0,\hat u,\hat v)(\bar x_0,\bar u,\bar v)\\ &=D\ell(\hat x_0,\hat x_T)(\bar x_0,\bar x_T)=-p_0\bar x_0+p_T\bar x_T\\ &=\int_0^T(\dot p\bar x+p\dot{\bar x})\mathrm{d}t=\int_0^T(D_uH\,\bar u+D_vH\,\bar v)\mathrm{d}t,\end{aligned}$$

where we used (9)-(10) and (12). Hence, necessarily one has

$$D_uH=0,\quad D_vH=0,\quad \text{a.e. on } [0,T].\tag{36}$$

This implies that $\lambda\in\Lambda$. On the other hand, simple computations yield that the second line of (35) is equivalent to

$$\Omega[\lambda](\bar x,\bar u,\bar v)\geq 0,\tag{37}$$

and, therefore, the result follows. $\square$

Remark 3.7. Observe that condition (27) can be extended to the cone $\mathcal{C}_2$ by the continuity of $\Omega[\lambda]$ and the compactness of Λ.

In the sequel we aim to strengthen previous necessary condition by proving that the maximum in (27) remains nonnegative when taken in a smaller set of multipliers. We shall first give a description of the subset of Lagrange multipliers we work with. Set

$$\mathcal{H}_2:=\{(\bar x,\bar u,\bar v)\in\mathcal{W}_2:(12)\text{ holds}\},\tag{38}$$

and consider the subset of Λ given by

$$\Lambda^\#:=\{\lambda\in\Lambda:\Omega[\lambda]\text{ is weakly-l.s.c. on }\mathcal{H}_2\}.$$

Lemma 3.8 below provides a characterization of $\Lambda^\#$ and Theorem 3.9 after gives a new necessary optimality condition. Recall first the definitions of R_0 and K given in (21).

Lemma 3.8.

$$\Lambda^\#=\{\lambda\in\Lambda:R_0[\lambda]\succeq 0\text{ and }K[\lambda]\equiv 0\}.\tag{39}$$

Theorem 3.9 (Strengthened second order necessary condition)**.** *If $\hat{w}$ is a weak minimum of problem (P), then*

$$\max_{\lambda\in\Lambda^{\#}} \Omega[\lambda](\bar{x},\bar{u},\bar{v}) \geq 0, \quad \text{on } \mathcal{C}_2. \tag{40}$$

In order to prove Lemma 3.8 notice that $\Omega[\lambda]$ can be written as the sum of two terms: the first one being a weakly-continuous function on the space $\mathcal{H}_2$ given by

$$(\bar{x},\bar{u},\bar{v}) \mapsto \tfrac{1}{2}D^2\ell[\lambda](\bar{x}_0,\bar{x}_T)^2 + \int_0^T [\tfrac{1}{2}\bar{x}^\top Q[\lambda]\bar{x} + \bar{u}^\top F[\lambda]\bar{x} + \bar{v}^\top C[\lambda]\bar{x}]\mathrm{d}t, \tag{41}$$

and the second one being the quadratic operator

$$(\bar{u},\bar{v}) \mapsto \int_0^T [\tfrac{1}{2}\bar{u}^\top R_0[\lambda]\bar{u} + \bar{v}^\top K[\lambda]\bar{u}]\mathrm{d}t. \tag{42}$$

The weak-continuity of the mapping in (41) is a consequence of Hestenes [32, Theorem 5.1]. On the other hand, in view of [32, Theorem 3.2] the following characterization holds.

Lemma 3.10. *The mapping in* (42) *is weakly-lower semicontinuous on* $\mathcal{U}\times\mathcal{V}$ *if and only if the matrix*

$$\begin{pmatrix} R_0[\lambda] & K[\lambda]^\top \\ K[\lambda] & 0 \end{pmatrix}, \tag{43}$$

is positive semidefinite a.e. on $[0,T]$.

Remark 3.11. The matrix in (43) is nothing but the second derivative of H with respect to the control (u,v). Therefore, the fact that this matrix is positive semidefinite is known as the *Legendre-Clebsch necessary optimality condition* for the extremal $(\hat{w},\lambda)$ (see e.g. [15, 2] or Corollary 3.13 below).

Notice now that Lemma 3.8 follows from the decomposition given by (41)-(42) and previous Lemma 3.10. On the other hand, Theorem 3.9 is a consequence of Remark 3.7, Lemma 3.8 and the following result on quadratic forms.

Lemma 3.12. [19, Theorem 5] *Given a Hilbert space H, and $a_1,a_2,\dots,a_p$ in H, set*

$$K := \{x\in H : (a_i,x)\leq 0, \text{ for } i=1,\dots,p\}. \tag{44}$$

Let M be a convex and compact subset of $\mathbb{R}^s$, and let $\{Q^\psi : \psi\in M\}$ be a family of continuous quadratic forms over H, the mapping $\psi\to Q^\psi$ being affine. Set $M^{\#} := \{\psi\in M :\ Q^\psi$ is weakly-l.s.c. on $H\}$ and assume that

$$\max_{\psi\in M} Q^\psi(x) \geq 0, \text{ for all } x\in K. \tag{45}$$

Then

$$\max_{\psi\in M^{\#}} Q^\psi(x) \geq 0, \text{ for all } x\in K. \tag{46}$$

We finish this section with the following Corollary.

Corollary 3.13 (Legendre-Clebsch condition)**.** *If $\hat{w}$ is a weak minimum of problem (P) with a unique associated multiplier $\hat{\lambda}$, then $(\hat{w}, \hat{\lambda})$ satisfies the Legendre-Clebsch condition. In order words, the matrix in* (43) *is positive semidefinite and, consequently,*

$$R_0[\hat{\lambda}] \succeq 0 \text{ and } K[\hat{\lambda}] \equiv 0. \tag{47}$$

Proof. It follows easily from Theorem 3.9. In fact, the inequality in (40) implies that $\Lambda^{\#} \neq \emptyset$, and since there is only one multiplier $\hat{\lambda}$, it follows that $\Lambda^{\#} = \{\hat{\lambda}\}$ and hence (47) necessarily holds. □

4. Goh Transformation

In this section we introduce a change of variables which consists of a linear transformation of $(\bar{x}, \bar{u}, \bar{v})$. The motivation of this change of variables is the following. In previous section we were able to provide a necessary condition involving the nonnegativity of the $\max \Omega[\lambda]$ on $\mathcal{C}_2$. The next step is to find a sufficient condition and, in order to achieve this, one would usually strengthen the inequality (40) to convert it into a condition of strong positivity. But since no quadratic term on $\bar{v}$ appears in Ω, the latter cannot be strongly positive. The technique we employ to find the desired sufficient condition is transforming Ω into a new quadratic mapping that may result strongly positive on an appropriate transformed critical cone. For historical interest, we recall that Goh introduced this change of variables in [28] and employed it to derive necessary conditions in [28, 26]. Afterwards, Dmitruk in [18] stated a second order sufficient condition for control-affine systems (case $l = 0$) in terms of the uniform positivity of $\max \Omega$ in the corresponding transformed space of variables.

Consider hence the linear system in (12) and the change of variables

$$\begin{cases} \bar{y}_t := \displaystyle\int_0^t \bar{v}_s \mathrm{d}s, \\ \bar{\xi}_t := \bar{x}_t - B_t \bar{y}_t, \end{cases} \quad \text{for } t \in [0, T]. \tag{48}$$

This change of variables can be done in any linear system of differential equations, and it is often called *Goh's transformation*. Observe that $\bar{\xi}$ defined in that way satisfies the linear equation

$$\dot{\bar{\xi}} = A\bar{\xi} + E\bar{u} + B_1 \bar{y}, \quad \bar{\xi}_0 = \bar{x}_0, \tag{49}$$

where A and E were given in (14), and

$$B_{1,t} := A_t B_t - \frac{\mathrm{d}}{\mathrm{d}t} B_t. \tag{50}$$

The $i-$th. column of B_1 is given by

$$\sum_{j=0}^m \hat{v}_j [f_j, f_i]^x + D_u f_i \, \dot{\hat{u}},$$

where $[f_i, f_j]^x := (D_x f_j) f_i - (D_x f_i) f_j$. Hence, we make the following hypothesis of regularity of the controls.

Assumption 4.1. The controls $\hat{u}$ and $\hat{v}$ are smooth.

In fact, a procedure of derivation of the controls as a function of the state and costate is done in Section 7 afterwards. It is proved that under Assumption 7.3, $(\hat{u}, \hat{v})$ can be written as a smooth function of $(\hat{x}, \lambda)$.

4.1. **Tranformed critical cones.** In this paragraph we present the critical cones obtained after Goh's transformation. Recall the linearized endpoint constraints in (16)-(17) and the critical cones given by (18)-(19). Let $(\bar{x}, \bar{u}, \bar{v}) \in \mathcal{C}$ be a critical direction. Define $(\bar{\xi}, \bar{y})$ by transformation (48) and set $\bar{h} := \bar{y}_T$. Note that (16)-(17) yield

$$D\eta_j(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{h}) = 0, \quad \text{for } j = 1, \ldots, d_\eta, \tag{51}$$

$$D\varphi_i(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{h}) \leq 0, \quad \text{for } i = 0, \ldots, d_\varphi. \tag{52}$$

Recall the definition of the linear space $\mathcal{W}_2$ in paragraph 2.3. Denote by $\mathcal{Y}$ the space $W^{1,\infty}(0,T;\mathbb{R}^m)$, and consider the cones

$$\mathcal{P} := \{(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \in \mathcal{W} \times \mathbb{R}^m : \bar{y}_0 = 0,\, \bar{y}_T = \bar{h},\, (49),\, (51)\text{-}(52) \text{ hold}\}, \tag{53}$$

$$\mathcal{P}_2 := \{(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \in \mathcal{W}_2 \times \mathbb{R}^m : \, (49),\, (51)\text{-}(52) \text{ hold}\}. \tag{54}$$

Remark 4.2. Notice that $\mathcal{P}$ consists of the directions obtained by transformating the elements of $\mathcal{C}$ via transformation (48).

The next result shows the density of $\mathcal{P}$ in $\mathcal{P}_2$. This fact is useful afterwards to extend a necessary condition in $\mathcal{P}$ to the bigger cone $\mathcal{P}_2$ by continuity arguments, as it was done for $\mathcal{C}$ and $\mathcal{C}_2$ in Section 3.

Lemma 4.3. *$\mathcal{P}$ is a dense subspace of $\mathcal{P}_2$ in the $\mathcal{W}_2 \times \mathbb{R}^m-$topology.*

Proof. Notice that the inclusion is immediate. In order to prove the density, consider the linear spaces

$$X := \{(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \in \mathcal{W}_2 \times \mathbb{R}^m : \, (49) \text{ holds}\}, \tag{55}$$

$$Y := \{(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \in \mathcal{W} \times \mathbb{R}^m : \, \bar{y}(0) = 0,\, \bar{y}(T) = \bar{h}, \text{ and } (49) \text{ holds}\}, \tag{56}$$

and the cone

$$Z := \{(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \in X : \, (51)\text{-}(52) \text{ holds}\}. \tag{57}$$

Notice that Y is a dense linear subspace of X (by Lemma 6 in [22] or Lemma 8.1 in [5]), and Z is a finite-faced cone of X. The desired density follows by Lemma 2.6. $\square$

4.2. **Transformed second variation.** Here we prove that performing the Goh's transformation in Ω yields the new quadratic operator $\Omega_{\mathcal{P}}$ in variables $(\bar{\xi}, \bar{u}, \bar{y}, \bar{v}, \bar{h})$ defined below and give a new necessary condition in terms of $\Omega_{\mathcal{P}}$. Recall the definitions in (21) and set, for $\lambda \in \Lambda_L^{\#}$,

$$\begin{aligned} \Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{v}, \bar{h}) &:= g[\lambda](\bar{\xi}_0, \bar{\xi}_T, \bar{h}) + \int_0^T \Big(\tfrac{1}{2}\bar{\xi}^\top Q[\lambda]\bar{\xi} + \bar{u}^\top F[\lambda]\bar{\xi} \\ &+ \bar{y}^\top M[\lambda]\bar{\xi} + \tfrac{1}{2}\bar{u}^\top R_0[\lambda]\bar{u} + \bar{y}^\top J[\lambda]\bar{u} + \tfrac{1}{2}\bar{y}^\top R_1[\lambda]\bar{y} + \bar{v}^\top V[\lambda]\bar{y} \Big)\,\mathrm{d}t, \end{aligned} \tag{58}$$

where

$$M := B^\top Q - \dot{C} - CA, \quad J := B^\top F^\top - CE, \tag{59}$$

$$S := \tfrac{1}{2}(CB + (CB)^\top), \quad V := \tfrac{1}{2}(CB - (CB)^\top), \tag{60}$$

$$R_1 := B^\top QB - (CB_1 + (CB_1)^\top) - \dot{S}, \tag{61}$$

$$g[\lambda](\zeta_0, \zeta_T, h) := \tfrac{1}{2}\ell''(\zeta_0, \zeta_T + B_T h)^2 + h^\top(C_T\zeta_T + \tfrac{1}{2}S_T h). \tag{62}$$

Observe that, in view of Assumptions 2.1 and 4.1, all the function defined above are continuous in time. Note that easy computations yield

$$V_{ij} = p[f_i, f_j]^x. \tag{63}$$

Theorem 4.4. *Let* $(\bar{x}, \bar{u}, \bar{v}) \in \mathcal{H}_2$ *(given in* (38)*) and* $(\bar{\xi}, \bar{y})$ *defined by the transformation* (48)*. Then*

$$\Omega[\lambda](\bar{x}, \bar{u}, \bar{v}) = \Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{v}, \bar{y}_T). \tag{64}$$

Proof. First recall that the term $\bar{v}^\top K[\lambda]\bar{u}$ in $\Omega[\lambda]$ vanishes since we are taking $\lambda \in \Lambda^\#$ and, in view of Lemma 3.8, $K[\lambda] \equiv 0$. In the remainder of the proof we omit the dependence on λ for the sake of simplicity. Replacing $\bar{x}$ in the definition of Ω in equation (20) by its expression in (48) yields
(65)

$$\begin{aligned}\Omega(\bar{x}, \bar{u}, \bar{v}) = \;& \tfrac{1}{2}\ell''(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{y}_T)^2 + \int_0^T \Big[\tfrac{1}{2}(\bar{\xi} + B\bar{y})^\top Q(\bar{\xi} + B\bar{y}) \\ & + \bar{u}^\top F(\bar{\xi} + B\bar{y}) + \bar{v}^\top C(\bar{\xi} + B\bar{y}) + \tfrac{1}{2}\bar{u}^\top R_0\bar{u}\Big]\,\mathrm{d}t.\end{aligned}$$

Integrate by parts the first term containing $\bar{v}$ in previous equation and use (49) to get

$$\int_0^T \bar{v}^\top C\bar{\xi}\mathrm{d}t = [\bar{y}^\top C\bar{\xi}]_0^T - \int_0^T \bar{y}^\top\{\dot{C}\bar{\xi} + C(A\bar{\xi} + E\bar{u} + B_1\bar{y})\}\mathrm{d}t. \tag{66}$$

The decomposition of CB introduced in (60) followed by an integration by parts leads to

$$\begin{aligned}\int_0^T \bar{v}^\top CB\bar{y}\mathrm{d}t &= \int_0^T \bar{v}^\top(S+V)\bar{y}\mathrm{d}t \\ &= \tfrac{1}{2}[\bar{y}^\top S\bar{y}]_0^T + \int_0^T(-\tfrac{1}{2}\bar{y}^\top\dot{S}\bar{y} + \bar{v}^\top V\bar{y})\mathrm{d}t.\end{aligned} \tag{67}$$

Combining (65), (66) and (67), the identity (64) follows. □

Finally recall Theorem 3.9. Observe that by performing Goh's transformation in (40) and in view of Remark 4.2, we obtain the following form of the second order necessary condition.

Corollary 4.5. *If* $\hat{w}$ *is a weak minimum of problem (P), then*

$$\max_{\lambda \in \Lambda^\#} \Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \dot{\bar{y}}, \bar{h}) \geq 0, \quad \text{on } \mathcal{P}. \tag{68}$$

5. New second order necessary condition

We aim to remove the dependence on $\bar v$ in the formulation of the second order necessary condition of Corollary 4.5. Note that in (68) $\bar v$ appears only in the term $\bar v^\top V[\lambda]\bar y$. Next we prove that we can restrict the maximum in (68) to the subset of $\Lambda_L^\#$ consisting of the multipliers for which $V[\lambda]$ vanishes.

Denote the convex hull of $\Lambda^\#$ by $\mathrm{co}\,\Lambda^\#$ and let $G(\mathrm{co}\,\Lambda^\#)$ be the subset of $\mathrm{co}\,\Lambda^\#$ for which $V[\lambda]$ vanishes, i.e.

$$G(\mathrm{co}\,\Lambda^\#) := \{\lambda \in \mathrm{co}\,\Lambda^\# : V[\lambda] \equiv 0\}. \tag{69}$$

The following optimality condition holds.

Theorem 5.1 (New necessary condition)**.** *If $\hat w$ is a weak minimum of problem (P), then*

$$\max_{\lambda \in G(\mathrm{co}\,\Lambda^\#)} \Omega_{\mathcal{P}}[\lambda](\bar\xi, \bar u, \bar y, \dot{\bar y}, \bar y_T) \geq 0, \quad \textit{on } \mathcal{P}. \tag{70}$$

Theorem 5.1 is an adaptation of very similar results given in Dmitruk [18] and Milyutin [43], that were employed recently in Aronna et al. [5]. The proof given in [5, Theorem 4.6] holds for Theorem 5.1 with minor modifications and hence we do not include it in the present article.

Notice that when $\hat w$ has a unique associated multiplier, from Theorem 5.1 we deduce that $G(\mathrm{co}\,\Lambda^\#)$ is not empty, and since the latter is a singleton, we get the corollary below. This corollary is one of the necessary conditions stated by Goh in [26].

Corollary 5.2. *Assume that $\hat w$ is a weak minimum having a unique associated multiplier. Then the following conditions holds.*

(i) $V \equiv 0$ *or, equivalently, the CB is symmetric or, in view of* (63),
$$p[f_i, f_j]^x = 0, \quad \textit{for } i, j = 1, \dots, m,$$
where p is the unique associated adjoint state.

(ii) *The matrix*
$$\begin{pmatrix} R_0 & J^\top \\ J & R_1 \end{pmatrix} \tag{71}$$
is positive semidefinite.

Observe that for $\lambda \in G(\mathrm{co}\,\Lambda^\#)$, the quadratic form $\Omega[\lambda]$ does not depend on $\bar v$ since its coefficients vanish. We can then consider its continuous extension to $\mathcal{P}_2$, given by

$$\begin{aligned} \Omega_{\mathcal{P}_2}[\lambda](\bar\xi, \bar u, \bar y, \bar h) := g[\lambda](\bar\xi_0, \bar\xi_T, \bar h) + \int_0^T \big(\tfrac{1}{2}\bar\xi^\top Q[\lambda]\bar\xi + \bar u^\top E[\lambda]\bar\xi \\ + \bar y^\top M[\lambda]\bar\xi + \tfrac{1}{2}\bar u^\top R_0[\lambda]\bar u + \bar y^\top J[\lambda]\bar u + \tfrac{1}{2}\bar y^\top R_1[\lambda]\bar y \big)\,\mathrm{d}t, \end{aligned} \tag{72}$$

where the involved matrices were defined in (14)-(15), (21), and (59)-(62). From Theorem 5.1 and previous definition, it follows:

Theorem 5.3. *If $\hat{w}$ is a weak minimum of problem (P), then*

$$\max_{\lambda \in G(\mathrm{co}\,\Lambda^{\#})} \Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \geq 0, \quad \text{on } \mathcal{P}_2. \tag{73}$$

Remark 5.4. The latter optimality condition does not involve $\bar{v}$. It is stated in the variable $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h})$.

6. Second order sufficient condition for weak minimum

This section provides a second order sufficient condition for strict weak optimality. Its proof is an adaptation of the proof of [5, Theorem 5.5] with important simplifications due to the absence of control constraints, but with some new difficulties owed to the presence of the nonlinear control variable.

Define the $\gamma-$*order* by

$$\gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{h}) := |\bar{x}_0|^2 + |\bar{h}|^2 + \int_0^T (|\bar{u}_t|^2 + |\bar{y}_t|^2)\mathrm{d}t, \tag{74}$$

for $(\bar{x}_0, \bar{u}, \bar{y}, \bar{h}) \in \mathbb{R}^n \times \mathcal{U}_2 \times \mathcal{V}_2 \times \mathbb{R}^m$. It can also be considered as a function of $(\bar{x}_0, \bar{u}, \bar{v}) \in \mathbb{R}^n \times \mathcal{U}_2 \times \mathcal{V}_2$ by setting

$$\tilde{\gamma}(\bar{x}_0, \bar{u}, \bar{v}) := \gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T), \tag{75}$$

with $\bar{y}$ being the primitive of $\bar{v}$ defined in (48).
Notation: We write γ to refer to either γ or $\tilde{\gamma}$.

Definition 6.1. [$\gamma-$growth] We say that $\hat{w}$ satisfies $\gamma-$*growth condition in the weak sense* if there exist $\varepsilon, \rho > 0$ such that

$$J(w) \geq J(\hat{w}) + \rho\gamma(x_0 - \hat{x}_0, u - \hat{u}, v - \hat{v}), \tag{76}$$

for every feasible trajectory w with $\|w - \hat{w}\|_\infty < \varepsilon$.

Theorem 6.2 (Sufficient condition for weak optimality)**.**

(i) *Assume that there exists $\rho > 0$ such that*

$$\max_{\lambda \in G(\mathrm{co}\,\Lambda^{\#})} \Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \geq \rho\gamma(\bar{\xi}_0, \bar{u}, \bar{y}, \bar{h}), \quad \text{on } \mathcal{P}_2. \tag{77}$$

Then $\hat{w}$ is a weak minimum satisfying $\gamma-$ growth in the weak sense.

(ii) *Conversely, if $\hat{w}$ is a weak solution satisfying $\gamma-$growth in the weak sense and such that $\alpha_0 > 0$ for every $\lambda \in G(\mathrm{co}\,\Lambda^{\#})$, then* (77) *holds for some $\rho > 0$.*

Corollary 6.3. *If $\hat{w}$ satisfies* (77) *and it has a unique associated multiplier, then necessarily the matrix in* (78) *is uniformly positive definite, i.e.*

$$\begin{pmatrix} R_0 & J^\top \\ J & R_1 \end{pmatrix} \succeq \rho I, \quad \text{on } [0, T], \tag{78}$$

where I refers to the identity matrix.

Remark 6.4. Another consequence of the condition (77) is stated in Remark 8.2 afterwards, where we link it with the *strengthened generalized Legendre-Clebsch condition.*

The remainder of this section is devoted to the proof of Theorem 6.2. We shall start by establishing some technical results that will be needed for the main result. For the lemma below recall the definition of the space $\mathcal{H}_2$ in (38).

Lemma 6.5. *There exists $\rho > 0$ such that*

$$|\bar{x}_0|^2 + \|\bar{x}\|_2^2 + |\bar{x}_T|^2 \leq \rho\gamma(\bar{x}_0, \bar{u}, \bar{v}), \tag{79}$$

for every linearized trajectory $(\bar{x}, \bar{u}, \bar{v}) \in \mathcal{H}_2$. The constant ρ depends on $\|A\|_\infty$, $\|B\|_\infty$, $\|E\|_\infty$ and $\|B_1\|_\infty$.

Proof. Throughout this proof, whenever we put ρ_i we refer to a positive constant depending on $\|A\|_\infty$, $\|B\|_\infty$, $\|E\|_\infty$, and/or $\|B_1\|_\infty$. Let $(\bar{x}, \bar{u}, \bar{v}) \in \mathcal{H}_2$ and $(\bar{\xi}, \bar{y})$ be defined by Goh's Transformation (48). Thus $(\bar{\xi}, \bar{u}, \bar{y})$ is solution of (49). Gronwall's Lemma and Cauchy-Schwartz inequality yield

$$\|\bar{\xi}\|_\infty \leq \rho_1(|\bar{\xi}_0|^2 + \|\bar{u}\|_2^2 + \|\bar{y}\|_2^2)^{1/2} \leq \rho_1\gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T)^{1/2}, \tag{80}$$

with $\rho_1 = \rho_1(\|A\|_1, \|E\|_\infty, \|B_1\|_\infty)$. This last inequality together with the relation between $\bar{\xi}$ and $\bar{x}$ provided by (48) imply

$$\|\bar{x}\|_2 \leq \|\bar{\xi}\|_2 + \|B\|_\infty\|\bar{y}\|_2 \leq \rho_2\gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T)^{1/2}, \tag{81}$$

for $\rho_2 = \rho_2(\rho_1, \|B\|_\infty)$. On the other hand, (48) and estimate (80) lead to

$$|\bar{x}_T| \leq |\bar{\xi}_T| + \|B\|_\infty|\bar{y}_T| \leq \rho_1\gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T)^{1/2} + \|B\|_\infty|\bar{y}_T|.$$

Then, in view of Young's inequality '$2ab \leq a^2 + b^2$' for real numbers a, b, one gets

$$|\bar{x}_T|^2 \leq \rho_3\gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T), \tag{82}$$

for some $\rho_3 = \rho_3(\rho_1, \|B\|_\infty)$. The desired estimate follows from (81) and (82). $\square$

Notice that Lemma 6.5 above gives an estimate of the linearized state in the order γ. The following result shows that the analogous property holds for the variation of the state variable as well.

Lemma 6.6. *Given $C > 0$, there exists $\rho > 0$ such that*

$$|\delta x_0|^2 + \|\delta x\|_2^2 + |\delta x_T|^2 \leq \rho\gamma(\delta x_0, \delta u, \delta v), \tag{83}$$

for every (x, u, v) solution of the state equation (2) having $\|v\|_2 \leq C$, and where $\delta w := w - \hat{w}$. The constant ρ depends on C, $\|B\|_\infty$, $\|\dot{B}\|_\infty$ and the Lipschitz constants of f_i.

Proof. In order to simplify the notation we omit the dependence on t. Consider (x, u, v) solution of (2) with $\|v\|_2 \leq C$. Let $\delta w := w - \hat{w}$, $\delta y := \int \delta v$,

and $\xi := \delta x - B\delta y$, with B given in (15) and $y_t := \int_0^t v_s \mathrm{d}s$. Note that

$$\begin{aligned}\dot{\xi} &= \sum_{i=0}^{m} [v_i f_i(x,u) - \hat{v}_i f_i(\hat{x},\hat{u})] - \dot{B}\delta y - \sum_{i=1}^{m} \delta v_i\, f_i(\hat{x},\hat{u}) \\ &= \sum_{i=0}^{m} v_i [f_i(x,u) - f_i(\hat{x},\hat{u})] - \dot{B}\delta y,\end{aligned} \tag{84}$$

where $v_0 \equiv 1$. In view of the Lipschitz-continuity of f_i,

$$|f_i(x,u) - f_i(\hat{x},\hat{u})| \le L(|\delta x| + |\delta u|) \le L(|\xi| + \|B\|_\infty |\delta y| + |\delta u|), \tag{85}$$

for some $L > 0$. Thus, from (84) it follows

$$\begin{aligned}&|\dot{\xi}| \le L(|\xi| + \|B\|_\infty |\delta y| + |\delta u|)(1+|v|) + \|\dot{B}\|_\infty |\delta y| \\ &= L\left\{|\xi|(1+|v|) + \|B\|_\infty |\delta y| + |\delta u| + \|B\|_\infty |\delta y||v| + |\delta u||v|\right\} + \|\dot{B}\|_\infty |\delta y|.\end{aligned}$$

Applying Gronwall's Lemma and Cauchy-Schwartz inequality to previous estimate yields

$$\|\xi\|_\infty \le \rho_1(|\xi_0| + \|\delta y\|_1 + \|\delta u\|_1 + \|\delta y\|_2 \|v\|_2 + \|\delta u\|_2 \|v\|_2), \tag{86}$$

for $\rho_1 = \rho_1(L, C, \|B\|_\infty, \|\dot{B}\|_\infty)$. Hence, since $\|\delta x\|_2 \le \|\xi\|_2 + \|B\|_\infty \|\delta y\|_2$, by previous estimate and Cauchy-Schwartz inequality, the desired result follows. □

Finally, the following lemma gives an estimate for the difference between the variation of the state variable and the linearized state.

Lemma 6.7. *Consider $C > 0$ and $w = (x,u,v) \in \mathcal{W}$ a feasible trajectory with $\|w - \hat{w}\|_\infty < C$. Set $(\delta x, \bar{u}, \bar{v}) := w - \hat{w}$ and let $\bar{x}$ be the linearization of $\hat{x}$ associated with $(\delta x_0, \bar{u}, \bar{v})$. Define*

$$\vartheta := \delta x - \bar{x}. \tag{87}$$

Then, ϑ is solution of the differential equation

$$\begin{aligned}&\dot{\vartheta} = \sum_{i=0}^{m} \hat{v}_i D_x f_i(\hat{x},\hat{u})\vartheta + \sum_{i=1}^{m} \bar{v}_i D f_i(\hat{x},\hat{u})(\delta x, \bar{u}) + \zeta, \\ &\vartheta_0 = 0,\end{aligned} \tag{88}$$

where the remainder ζ satisfies the estimates

$$\|\zeta\|_\infty < \rho_1, \quad \|\zeta\|_2 < \rho_2\, \gamma, \tag{89}$$

where ρ_1, ρ_2 depend on C, $\|D^2 f\|_\infty$ and the Lipschitz constant of $D^2 f$. If in addition, $\|\bar{u}\|_2 + \|\bar{v}\|_2 \to 0$, the following estimates for ϑ hold

$$\|\vartheta\|_\infty < o(\sqrt{\gamma}), \quad \|\dot{\vartheta}\|_2 < o(\sqrt{\gamma}). \tag{90}$$

Proof. Let us begin by observing that the variation of the state variable satisfies the differential equation

$$\dot{\delta x} = \sum_{i=0}^{m} v_i [f_i(x,u) - f_i(\hat{x},\hat{u})] + \sum_{i=1}^{m} \bar{v}_i f_i(\hat{x},\hat{u}). \tag{91}$$

Consider the following Taylor expansions for f_i,

$$f_i(x,u) = f_i(\hat{x},\hat{u}) + Df_i(\hat{x},\hat{u})(\delta x,\bar{u}) + \tfrac{1}{2}D^2 f_i(\hat{x},\hat{u})(\delta x,\bar{u})^2 + \rho_0|(\delta x,\bar{u})|^3, \tag{92}$$

where ρ_0 is a function of the Lipschitz constant of $D^2 f_i$. Combining (91) and (92) yields

$$\dot{\delta x} = \sum_{i=0}^m v_i Df_i(\hat{x},\hat{u})(\delta x,\bar{u}) + \sum_{i=1}^m \bar{v}_i f_i(\hat{x},\hat{u}) + \zeta, \tag{93}$$

with the remainder being given by

$$\zeta := \tfrac{1}{2}\sum_{i=0}^m v_i [D^2 f_i(\hat{x},\hat{u})(\delta x,\bar{u})^2 + \rho_0|(\delta x,\bar{u})|^3. \tag{94}$$

The linearized equation (12) together with (93) lead to (88), and, in view of (94), it can be seen that the estimates in (89) hold. Applying Gronwall's Lemma in (88), and using Cauchy-Schwartz inequality afterwards lead to

$$\|\vartheta\|_\infty \leq \rho_3 \left\| \sum_{i=1}^m \bar{v}_i Df_i(\hat{x},\hat{u})(\delta x,\bar{u}) + \zeta \right\|_1 \leq \rho_4 \left[\|\bar{v}\|_2(\|\delta x\|_2 + \|\bar{u}\|_2) + \|\zeta\|_2 \right],$$

for some positive ρ_3, ρ_4 depending on C and $\|Df\|_\infty$. Finally, using the estimate in Lemma 6.6 and (89) just obtained, the inequalities in (90) follow. $\square$

In view of Lemmas 3.1, 6.5, 6.6 and 6.7 we can justify the following technical result that is an essential point in the proof of the sufficient condition of Theorem 6.2.

Lemma 6.8. *Let $w \in \mathcal{W}$ be a feasible solution. Set $(\delta x, \bar{u}, \bar{v}) := w - \hat{w}$, and $\bar{x}$ its corresponding linearized state, i.e. the solution of* (12)-(13) *associated with $(\delta x_0, \bar{u}, \bar{v})$. Assume that $\|w - \hat{w}\|_\infty \to 0$. Then*

$$Ł[\lambda](w) = Ł[\lambda](\hat{w}) + \Omega[\lambda](\bar{x},\bar{u},\bar{v}) + o(\gamma). \tag{95}$$

Proof. Omit the dependence on λ for the sake of simplicity. Recall the expansion of the Lagrangian function given in Lemma 3.1. Notice that by Lemma 6.6, $Ł(w) = Ł(\hat{w}) + \Omega(\delta x,\bar{u},\bar{v}) + o(\gamma)$. Hence,

$$Ł(w) = Ł(\hat{w}) + \Omega(\bar{x},\bar{u},\bar{v}) + \Delta\Omega + o(\gamma), \tag{96}$$

with $\Delta\Omega := \Omega(\delta x,\bar{u},\bar{v}) - \Omega(\bar{x},\bar{u},\bar{v})$. The next step is using Lemmas 6.5, 6.6 and 6.7 to prove that

$$\Delta\Omega = o(\gamma). \tag{97}$$

Note that $\mathcal{Q}(a,a) - \mathcal{Q}(b,b) = \mathcal{Q}(a+b,a-b)$, for any bilinear mapping $\mathcal{Q}$, and any pair a,b of elements in its domain. Set $\vartheta := \delta x - \bar{x}$ as it is done in Lemma 6.7. Hence,

$$\Delta\Omega = \tfrac{1}{2}\ell''((\delta x_0+\bar{x}_0, \delta x_T+\bar{x}_T),(0,\vartheta_T)) + \int_0^T [\tfrac{1}{2}(\delta x+\bar{x})^\top Q\vartheta + \bar{u}^\top E\vartheta + \bar{v}^\top C\vartheta]\mathrm{d}t.$$

The estimates in Lemmas 6.5 and 6.6 yield $\Delta\Omega = \int_0^T \bar{v}^\top C\vartheta \mathrm{d}t + o(\gamma)$. Integrating by parts in the latter expression and using (90) lead to

$$\int_0^T \bar{v}^\top C\vartheta \mathrm{d}t = [\bar{y}^\top C\vartheta]_0^T - \int_0^T \bar{y}^\top (\dot{C}\vartheta + C\dot{\vartheta})\mathrm{d}t = o(\gamma),$$

and hence the desired result follows. □

Proof. [of Theorem 6.2] We shall prove that if (77) holds for some $\rho > 0$, then $\hat{w}$ satisfies $\gamma-$growth in the weak sense. By the contrary assume that the $\gamma-$growth condition (76) is not satisfied. Consequently, there exists a sequence of feasible trajectories $\{w_k\}$ converging to $\hat{w}$ in the weak sense, such that

$$J(w_k) \leq J(\hat{w}) + o(\gamma_k), \tag{98}$$

with $\delta w_k := w_k - \hat{w}$ and $\gamma_k := \gamma(\delta x_{k,0}, \bar{u}_k, \bar{v}_k)$. Let $(\bar{\xi}_k, \bar{u}_k, \bar{y}_k)$ be the transformed directions defined by (48). We divide the remainder of the proof in two steps.

(I) First we prove that the sequence given by

$$(\tilde{\xi}_k, \tilde{u}_k, \tilde{y}_k, \tilde{h}_k) := (\bar{\xi}_k, \bar{u}_k, \bar{y}_k, \bar{h}_k)/\sqrt{\gamma_k} \tag{99}$$

contains a subsequence converging to an element $(\tilde{\xi}, \tilde{u}, \tilde{y}, \tilde{h})$ of $\mathcal{P}_2$ in the weak topology, i.e. $(\tilde{u}_k, \tilde{y}_k) \rightharpoonup (\tilde{u}, \tilde{y})$ in the weak topology of $\mathcal{U}_2 \times \mathcal{V}_2$ and $(\tilde{\xi}_k, \tilde{h}_k) \to (\tilde{\xi}, \tilde{h})$ in the strong sense of $\mathcal{X}_2 \times \mathbb{R}^m$.

(II) Afterwards, employing the latter sequence and its weak limit, we show that (77) together with (98) lead to a contradiction.

We shall begin by Part (I). For this we take an arbitrary Lagrange multiplier λ in $\Lambda_L^\#$. By multiplying the inequality (98) by α_0, and adding the nonpositive term

$$\sum_{i=0}^{d_\varphi} \alpha_i \varphi_i(x_{k,0}, x_{k,T}) + \sum_{j=1}^{d_\eta} \beta_j \eta_j(x_{k,0}, x_{k,T}), \tag{100}$$

to its left-hand side, the inequality follows

$$Ł[\lambda](w_k) \leq Ł[\lambda](\hat{w}) + o(\gamma_k). \tag{101}$$

Let us now recall the expansion (95) given in Lemma 6.8. Note that the elements of the sequence $(\tilde{\xi}_{k,0}, \tilde{u}_k, \tilde{y}_k, \tilde{h}_k)$ have unit $\mathbb{R}^n \times \mathcal{U}_2 \times \mathcal{V}_2 \times \mathbb{R}^m-$norm. The Banach-Alaoglu Theorem (see e.g. [14, Theorem III.15]) implies that, extracting if necessary a subsequence, there exists $(\tilde{\xi}_0, \tilde{u}, \tilde{y}, \tilde{h}) \in \mathbb{R}^n \times \mathcal{U}_2 \times \mathcal{V}_2 \times \mathbb{R}^m$ such that

$$\tilde{\xi}_{k,0} \to \tilde{\xi}_0, \quad \tilde{u}_k \rightharpoonup \tilde{u}, \quad \tilde{y}_k \rightharpoonup \tilde{y}, \quad \tilde{h}_k \to \tilde{h}, \tag{102}$$

where the two limits indicated with $\rightharpoonup$ are taken in the weak topology of $\mathcal{U}_2$ and $\mathcal{V}_2$, respectively. The solution of equation (49) associated with $(\tilde{\xi}_0, \tilde{u}, \tilde{y})$ is denoted by $\tilde{\xi}$, which is the limit of $\tilde{\xi}_k$ in $\mathcal{X}_2$. For the aim of proving

that $(\tilde{\xi},\tilde{u},\tilde{v},\tilde{h})$ belongs to $\mathcal{P}_2$, we shall check that the initial-final conditions (51)-(52) are verified. For each index $0\leq i\leq d_\varphi$, one has

$$D\varphi_i(\hat{x}_0,\hat{x}_T)(\tilde{\xi}_0,\tilde{\xi}_T+B_T\tilde{h})=\lim_{k\to\infty}D\varphi_i(\hat{x}_0,\hat{x}_T)\left(\frac{\bar{x}_{k,0},\bar{x}_{k,T}}{\sqrt{\gamma_k}}\right). \tag{103}$$

In order to prove that the right hand-side of (103) is nonpositive, we consider the following first order Taylor expansion of function φ_i around $(\hat{x}_0,\hat{x}_T)$:

$$\varphi_i(x_{k,0},x_{k,T})=\varphi_i(\hat{x}_0,\hat{x}_T)+D\varphi_i(\hat{x}_0,\hat{x}_T)(\delta x_{k,0},\delta x_{k,T})+o(|(\delta x_{k,0},\delta x_{k,T})|).$$

Previous equation and Lemmas 6.5 and 6.7 imply

$$\varphi_i(x_{k,0},x_{k,T})=\varphi_i(\hat{x}_0,\hat{x}_T)+D\varphi_i(\hat{x}_0,\hat{x}_T)(\bar{x}_{k,0},\bar{x}_{k,T})+o(\sqrt{\gamma_k}).$$

Thus, the following approximation for the right hand-side in (103) holds,

$$D\varphi_i(\hat{x}_0,\hat{x}_T)\left(\frac{\bar{x}_{k,0},\bar{x}_{k,T}}{\sqrt{\gamma_k}}\right)=\frac{\varphi_i(x_{k,0},x_{k,T})-\varphi_i(\hat{x}_0,\hat{x}_T)}{\sqrt{\gamma_k}}+o(1). \tag{104}$$

Since w_k is a feasible trajectory, it satisfies (4) and, therefore, equations (103) and (104) yield, for $1\leq i\leq d_\varphi$, $D\varphi_i(\hat{x}_0,\hat{x}_T)(\tilde{\xi}_0,\tilde{\xi}_T+B_T\tilde{h})\leq 0$. For $i=0$ use inequality (98) to get the corresponding inequality. Analogously,

$$D\eta_j(\hat{x}_0,\hat{x}_T)(\tilde{\xi}_0,\tilde{\xi}_T+B_T\tilde{h})=0,\quad \text{for } j=1,\dots,d_\eta. \tag{105}$$

Thus $(\tilde{\xi},\tilde{u},\tilde{y},\tilde{h})$ satisfies (51)-(52), and hence it belongs to $\mathcal{P}_2$.

Let us deal with Part (II). Notice that from (95) and (101) we get

$$\Omega_{\mathcal{P}_2}[\lambda](\tilde{\xi}_k,\tilde{u}_k,\tilde{y}_k,\tilde{h}_k)\leq o(1), \tag{106}$$

and thus

$$\liminf_{k\to\infty}\Omega_{\mathcal{P}_2}[\lambda](\tilde{\xi}_k,\tilde{u}_k,\tilde{y}_k,\tilde{h}_k)\leq 0. \tag{107}$$

Consider the subset of $G(\operatorname{co}\Lambda_L^{\#})$ defined by

$$\Lambda_L^{\#,\rho}:=\{\lambda\in G(\operatorname{co}\Lambda_L^{\#}):\Omega_{\mathcal{P}_2}[\lambda]-\rho\gamma \text{ is weakly l.s.c. on } \mathcal{H}_2\times\mathbb{R}^m\}. \tag{108}$$

By applying Lemma 3.12 to the inequality (77) one has

$$\max_{\lambda\in\Lambda_L^{\#,\rho}}\Omega_{\mathcal{P}_2}[\lambda](\bar{\xi},\bar{u},\bar{y},\bar{h})\geq\rho\gamma(\bar{\xi}_0,\bar{u},\bar{y},\bar{h}),\quad \text{on } \mathcal{P}_2. \tag{109}$$

We shall take $\tilde{\lambda}\in\Lambda_L^{\#,\rho}$ that attains the maximum in (109) for the direction $(\tilde{\xi},\tilde{u},\tilde{y},\tilde{h})$. Hence we get

$$\begin{aligned}0&\leq\Omega_{\mathcal{P}_2}[\tilde{\lambda}](\tilde{\xi},\tilde{u},\tilde{y},\tilde{h})-\rho\gamma(\tilde{\xi}_0,\tilde{u},\tilde{y},\tilde{h})\\&\leq\liminf_{k\to\infty}\Omega_{\mathcal{P}_2}[\tilde{\lambda}](\tilde{\xi}_k,\tilde{u}_k,\tilde{y}_k,\tilde{h}_k)-\rho\gamma(\tilde{\xi}_{k,0},\tilde{u}_k,\tilde{y}_k,\tilde{h}_k)\leq-\rho,\end{aligned} \tag{110}$$

since $\Omega_{\mathcal{P}_2}[\tilde{\lambda}]-\rho\gamma$ is weakly-l.s.c., $\gamma(\tilde{\xi}_{k,0},\tilde{u}_k,\tilde{y}_k,\tilde{h}_k)=1$ for every k and inequality (107) holds. This leads us to a contradiction since $\rho>0$. Therefore, the desired result follows.

(ii) Let us now prove the second statement. Assume that $\hat{w}$ is a weak solution satisfying $\gamma-$growth in the weak sense for some constant $\rho'>0$,

and such that $\alpha_0 > 0$ for every multiplier $\lambda \in G(\operatorname{co}\Lambda^\#)$. We consider the modified problem

$$(\tilde{P}) \qquad \begin{aligned} & J(w) - \rho' \gamma(w - \hat{w}) \to \min, \\ & \text{s.t. } (2)\text{-}(4) \end{aligned}$$

and rewrite it in the Mayer form

$$(\tilde{\tilde{P}}) \qquad \begin{aligned} & J(w) - \rho'(|x_0 - \hat{x}_0|^2 + |y_T - \hat{y}_T|^2 + \pi_{1,T} + \pi_{2,T}) \to \min, \\ & \text{s.t. } (2) - (4), \\ & \dot{y} = v, \\ & \dot{\pi}_1 = (u - \hat{u})^2, \\ & \dot{\pi}_2 = (y - \hat{y})^2, \\ & y_0 = 0, \ \pi_{1,0} = 0, \ \pi_{2,0} = 0. \end{aligned}$$

We aim to apply the second order necessary condition of Theorem 5.3 to $(\tilde{\tilde{P}})$ at the point $(w = \hat{w}, y = \hat{y}, \pi_1 = 0, \pi_2 = 0)$. Simple computations show that at this solution each critical cone of (53) is the projection of the corresponding critical cone of $(\tilde{\tilde{P}})$, and that the same holds for the set of multipliers. Furthermore, the second variation of $(\tilde{\tilde{P}})$ evaluated at a multiplier $\tilde{\tilde{\lambda}} \in G(\operatorname{co}\tilde{\tilde{\Lambda}}^\#)$ is given by

$$\Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{y}_T) - \alpha_0 \rho' \gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T), \tag{111}$$

where $\lambda \in G(\operatorname{co}\Lambda^\#)$ is the corresponding multiplier for problem (53). Hence, the necessary condition in Theorem 5.3 implies that for every $(\bar{\xi}, \bar{u}, \bar{v}, \bar{h}) \in \mathcal{P}_2$ there exists $\lambda \in G(\operatorname{co}\Lambda^\#)$ such that

$$\Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}, \bar{u}, \bar{y}, \bar{y}_T) - \alpha_0 \rho' \gamma(\bar{x}_0, \bar{u}, \bar{y}, \bar{y}_T) \geq 0.$$

Setting $\rho := \min_{G(\operatorname{co}\Lambda^\#)} \alpha_0 \rho' > 0$ yields the desired result. This completes the proof of the theorem. □

7. Shooting algorithm

The purpose of this section is to present an appropriate numerical scheme to solve the problem given by equations (1)-(3), which we denote (SP). Notice that no inequality endpoint constraints are considered. More precisely, we investigate the formulation and the convergence of an algorithm that approximates an optimal solution provided an initial estimate.

We shall consider an hypothesis concerning the endpoint conditions. With this end recall Definition 3.4. The following holds throughout the rest of the article.

Assumption 7.1. The endpoint equality constraints are qualified or, equivalently, the derivative of $\bar{\eta}$ at $(\hat{x}_0, \hat{u}, \hat{v})$ is onto.

It is a well-known result that in this case $\hat{w}$ is normal and has a unique associated multiplier (see e.g. Pontryagin et al. [52]). Therefore, without loss of generality, we can consider $\alpha_0 = 1$. The unique multiplier associated with $\hat{w}$ is denoted by $\hat{\lambda} = (\hat{\beta}, \hat{p})$.

7.1. **Optimality system.** In what follows we use the first order optimality conditions (11) to provide a set of equations from which we can determine $\hat{w}$. We obtain an optimality system in the form of a *two-point boundary value problem* (TPBVP).

We shall recall that for the case where all the control variables appear nonlinearly ($m = 0$), the classical technique is using the stationarity equation

$$H_u[\hat{\lambda}](\hat{w}) = 0, \tag{112}$$

to write $\hat{u}$ as a function of $(\hat{x}, \hat{\lambda})$ (this is done in e.g. [16, 40, 12, 54]). One is able to do this by assuming, for instance, the *strengthened Legendre-Clebsch condition*

$$H_{uu}[\hat{\lambda}](\hat{w}) \succ 0. \tag{113}$$

The latter condition comes from strengthening the inequality in the necessary optimality condition mentioned in Remark 3.11, which is verified by $\hat{w}$ in view of Corollary 3.13. In this case, due to the Implicit Function Theorem, we can write $\hat{u} = U[\hat{\lambda}](\hat{x})$ with U being a smooth function. Hence, replacing the occurrences of $\hat{u}$ by $U[\hat{\lambda}](\hat{x})$ in the state and costate equations yields a two-point boundary value problem.

On the other hand, when the system is affine in all the control variables ($l = 0$), we cannot eliminate the control from the equation $H_v = 0$ and, therefore, a different technique is employed (see e.g. [39, 47, 51, 12, 7, 54]). The idea is to consider an index $1 \leq i \leq m$, and to take $d^{M_i} H_v / dt^{M_i}$ to be the lowest order derivative of H_v in which $\hat{v}_i$ appears with a coefficient that is not identically zero. Kelley [34], Goh [28, 27], Kelley et al. [35] and Robbins [53] proved that M_i is even when the investigated extremal is normal. This implies that $\dot{H}_v$ depends only on $\hat{x}$ and $\hat{\lambda}$ and, consequently, it is differentiable in time. Thus the expression

$$\ddot{H}_v[\hat{\lambda}](\hat{w}) = 0 \tag{114}$$

is well-defined. The control $\hat{v}$ can be retrieved from (114) provided that, for instance, the *strengthened generalized Legendre-Clebsch condition*

$$-\frac{\partial \ddot{H}_v}{\partial v}[\hat{\lambda}](\hat{w}) \succ 0 \tag{115}$$

holds (see Goh [27, 30, 31]). In this case, we can write $\hat{v} = V[\hat{\lambda}](\hat{x})$ with V being differentiable. By replacing $\hat{v}$ by $V[\hat{\lambda}](\hat{x})$ in the state-costate equations, we get an optimality system in the form of a boundary value problem.

In the problem studied here, where $l > 0$ and $m > 0$, we aim to use both equations (112) and (114) to retrieve the control $(\hat{u}, \hat{v})$ as a function of the

state $\hat{x}$ and the multiplier $\hat{\lambda}$. We next describe a procedure to achieve this elimination that was proposed in Goh [30, 31]. Let us show that H_v can be differentiated two times in the time variable, as it was done in the totally affine case. Observe that (112) may be used to write $\dot{\hat{u}}$ as a function of $(\hat{\lambda}, \hat{w})$. In fact, in view of Corollary 3.13,

$$H_{uv} = 0, \tag{116}$$

and hence the coefficient of $\dot{\hat{v}}$ in $\dot{H}_u$ is zero. Consequently,

$$\dot{H}_u = \dot{H}_u[\hat{\lambda}](\hat{x}, \hat{u}, \hat{v}, \dot{\hat{u}}) = 0 \tag{117}$$

and, if the strengthened Legendre-Clebsch condition (113) holds, $\dot{\hat{u}}$ can be eliminated from (117) yielding

$$\dot{\hat{u}} = \Gamma[\hat{\lambda}](\hat{x}, \hat{u}, \hat{v}). \tag{118}$$

Take now an index $i = 1, \ldots, m$ and observe that

$$0 = \dot{H}_{v_i} = \frac{\mathrm{d}}{\mathrm{d}t}\hat{p}\hat{f}_i = \hat{p}\sum_{j=0}^{m} \hat{v}_j [f_j, f_i]^x(\hat{x}, \hat{u}) + H_{v_i u}\dot{\hat{u}} = \hat{p}\,[f_0, f_i]^x(\hat{x}, \hat{u}), \tag{119}$$

where Corollary 5.2 and (116) are used in the last equality. Therefore, $\dot{H}_v = \dot{H}_v[\hat{\lambda}](\hat{x}, \hat{u})$. We can then differentiate one more time $\dot{H}_v$, replace the occurrence of $\dot{\hat{u}}$ by Γ in (118) and obtain (114) as it was desired. See that (114) together with the boundary conditions

$$H_v[\hat{\lambda}](\hat{w}_T) = 0, \tag{120}$$

$$\dot{H}_v[\hat{\lambda}](\hat{w}_0) = 0, \tag{121}$$

guarantee the second identity in the stationarity condition (11).
Notation: Denote by (OS) the set of equations consisting of (2)-(3), (7), (9)-(10), (112), (114) and the boundary conditions (120)-(121).

Remark 7.2. Instead of (120)-(121), we could choose another pair of end-point conditions among the four possible ones: $H_{v,0} = 0$, $H_{v,T} = 0$, $\dot{H}_{v,0} = 0$ and $\dot{H}_{v,T} = 0$, always including at least one of order zero. The choice we made will simplify the presentation of the result afterwards.

Observe now that the derivative of the mapping $(u, v) \mapsto \begin{pmatrix} H_u \\ -\ddot{H}_v \end{pmatrix}$ is given by

$$\mathcal{J} := \begin{pmatrix} H_{uu} & H_{uv} \\ -\dfrac{\partial \ddot{H}_v}{\partial u} & -\dfrac{\partial \ddot{H}_v}{\partial v} \end{pmatrix}. \tag{122}$$

On the other hand, if (113) and (115) are verified, $\mathcal{J}$ is definite positive and consequently, nonsingular. In this case we may write $\hat{u} = U[\hat{\lambda}](\hat{x})$ and $\hat{v} = V[\hat{\lambda}](\hat{x})$ from (112) and (114). Thus (OS) can be regarded as a TPBVP whenever the following hypothesis is verified.

Assumption 7.3. The conditions (113) and (115) hold along $\hat{w}$.

Summing up we get the following result.

Proposition 7.4 (Elimination of the control). *If Assumption 7.3 holds, then one has*

$$\hat{u} = U[\hat{\lambda}](\hat{x}), \quad \hat{v} = V[\hat{\lambda}](\hat{x}),$$

for smooth functions U and V.

Remark 7.5. When the linear and nonlinear controls are uncoupled, this elimination of the controls is much simpler. An example is shown in Oberle [48] where a nonlinear control variable can be eliminated by the stationarity of the pre-Hamiltonian, and the remaining problem has two uncoupled controls, one linear and one nonlinear.

The rest of this article is very close to what was done in Aronna et al. [7]. The main difference between the totally affine case and the mixed case treated here lies on the derivation of the system (OS). The proof of the convergence in Section 8 is an extension of the proof of Theorem 5 in [7]. The presentation here is then more concise, and the reader is referred to the mentioned article for further details.

7.2. **The algorithm.** The aim of this section is to present a numerical scheme to solve system (OS). In view of Proposition 7.4 we can define the following mapping.

Definition 7.6. Let $\mathcal{S} : \mathbb{R}^n \times \mathbb{R}^{n+d_\eta,*} =: \mathrm{D}(\mathcal{S}) \to \mathbb{R}^{d_\eta} \times \mathbb{R}^{2n+2m,*}$ be the *shooting function* given by

$$(x_0, p_0, \beta) =: \nu \quad \mapsto \quad \mathcal{S}(\nu) := \begin{pmatrix} \eta(x_0, x_T) \\ p_0 + D_{x_0}\ell[\lambda](x_0, x_T) \\ p_T - D_{x_T}\ell[\lambda](x_0, x_T) \\ H_v[\lambda](w_T) \\ \dot{H}_v(w_0) \end{pmatrix}, \tag{123}$$

where (x, p) is a solution of (2),(9),(112),(114) with initial conditions x_0 and p_0, and $\lambda := (p, \beta)$, and where the occurrences of u and v were replaced by $u = U[\lambda](x)$ and $v = V[\lambda](x)$.

Note that solving (OS) consists of finding $\hat{\nu} \in \mathrm{D}(\mathcal{S})$ such that

$$\mathcal{S}(\hat{\nu}) = 0. \tag{124}$$

Since the number of equations in (124) is greater than the number of unknowns, the Gauss-Newton method is a suitable approach to solve it. The *shooting algorithm* we propose here consists of solving the equation (124) by the Gauss-Newton method. A more extensive description of this algorithm is presented in [7]. There it is observed that the method is applicable provided that $\mathcal{S}'(\hat{\nu})$ is one-to-one, with $\hat{\nu} := (\hat{x}_0, \hat{p}_0, \hat{\beta})$. Furthermore, since the right hand-side of system (124) is zero, it converges locally quadratically if the function $\mathcal{S}$ has Lipschitz continuous derivative. The latter holds true

here given the regularity hypotheses on the data functions (in Assumption 2.1). This convergence result is stated in the proposition below. See e.g. Fletcher [24] for a proof.

Proposition 7.7. *If $\mathcal{S}'(\hat{\nu})$ is one-to-one then the shooting algorithm is locally quadratically convergent.*

8. Convergence of the shooting algorithm: application of the second order sufficient condition

The main result of this last part of the article is the theorem below that gives a condition guaranteeing the quadratic convergence of the shooting method near an optimal local solution.

Theorem 8.1. *Suppose that $\hat{w}$ is such that* (77) *holds. Then the shooting algorithm is locally quadratically convergent.*

The idea is linking the sufficient condition (77) to the derivative $\mathcal{S}'(\hat{\nu})$. Notice that (77) is expressed in the variables after Goh's Transformation, while $\mathcal{S}$ is in the original variables. The procedure to achieve Theorem 8.1 has three stages that are described in the paragraphs 8.1, 8.2 and 8.3 below. The proof of 8.1 is at the end of 8.3.

Remark 8.2. In view of a result in [27, Section 4.8] the positive definiteness in (78) implies both (113) and (115). Therefore, in Theorem 8.1, the Assumption 7.3 is guaranteed by the condition (77).

8.1. **Linearization of (OS).** We write the linearized system associated with (OS), which gives the derivative of $\mathcal{S}$. A definition of linearized differential algebraic system can be found in e.g. Kunkel-Mehrmann [36] or Aronna et al. [7]. We denote by $\operatorname{Lin}\mathcal{F}$ the *linearization* of function $\mathcal{F}$, i.e.

$$\operatorname{Lin}\mathcal{F}\,|_{(\zeta_t^0,\alpha_t^0)}\,(\bar{\zeta}_t,\bar{\alpha}_t) := \mathcal{F}'(\zeta_t^0,\alpha_t^0)(\bar{\zeta}_t,\bar{\alpha}_t), \tag{125}$$

The technical result below will simplify the computation afterwards. Its proof is immediate (or see [36]).

Lemma 8.3 (Commutation of linearization and differentiation)**.** *Given $\mathcal{G}$ and $\mathcal{F}$ as in the previous definition, it holds:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Lin}\mathcal{G} = \operatorname{Lin}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}, \quad \frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Lin}\mathcal{F} = \operatorname{Lin}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}. \tag{126}$$

Recall the definitions in (14), (15) and (21). Notice that, since $H_v = pB$,

$$\operatorname{Lin} H_v = \bar{p}B + \bar{x}^\top C^\top. \tag{127}$$

Here whenever the argument of a function is missing, assume that it is evaluated on $(\hat{w},\hat{\lambda})$. The linearization of system (OS) at point $(\hat{x},\hat{u},\hat{v},\hat{\lambda})$ consists of the linearized state equation (12) with endpoint condition (16), the linearized costate equation

$$-\dot{\bar{p}}_t = \bar{p}_t A_t + \bar{x}_t^\top Q_t + \bar{u}_t^\top E_t + \bar{v}_t^\top C_t, \quad \text{a.e. on } [0,T], \tag{128}$$

with boundary conditions

$$\bar{p}_0 = -\left[\bar{x}_0^\top D^2_{x_0^2}\ell + \bar{x}_T^\top D^2_{x_0 x_T}\ell + \sum_{j=1}^{d_\eta} \bar{\beta}_j D_{x_0}\eta_j\right]_{(\hat{x}_0,\hat{x}_T)}, \tag{129}$$

$$\bar{p}_T = \left[\bar{x}_T^\top D^2_{x_T^2}\ell + \bar{x}_0^\top D^2_{x_0 x_T}\ell + \sum_{j=1}^{d_\eta} \bar{\beta}_j D_{x_T}\eta_j\right]_{(\hat{x}_0,\hat{x}_T)}, \tag{130}$$

and the algebraic equations

$$0 = \text{Lin } H_u = \bar{p}D + \bar{x}^\top E^\top + \bar{u}^\top R_0, \tag{131}$$

$$0 = \text{Lin } \ddot{H}_v = -\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\bar{p}B + \bar{x}^\top C^\top), \quad \text{a.e. on } [0,T], \tag{132}$$

$$0 = (\text{Lin } H_v)_T = \bar{p}_T B_T + \bar{x}_T^\top C_T^\top, \tag{133}$$

$$0 = (\text{Lin } \dot{H}_v)_0 = -\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(\bar{p}B + \bar{x}^\top C^\top), \tag{134}$$

where we used equation (127) and the commutation property of Lemma 8.3. Observe that (132) -(134) and Lemma 8.3 yield

$$0 = \text{Lin } H_v = \bar{p}B + \bar{x}^\top C^\top, \quad \text{a.e. on } [0,T]. \tag{135}$$

Notation: denote by (LS) the set of equations consisting of (12), (16), (128)-(134).

Proposition 8.4. *The differential $\mathcal{S}'(\hat{\nu})$ is one-to-one if the only solution of* (12), (128), (131), (132) *with the initial conditions $(\bar{x}_0, \bar{p}_0) = 0$ and with $\bar{\beta} = 0$ is $(\bar{x}, \bar{u}, \bar{v}, \bar{p}) = 0$.*

8.2. **Auxiliary linear-quadratic problem.** Now we introduce the following linear-quadratic control problem in the variables $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h})$. Denote by (LQ) the problem given by

$$\Omega_{\mathcal{P}_2}(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \to \min, \tag{136}$$

$$(49),(51), \tag{137}$$

$$\dot{\bar{h}} = 0. \tag{138}$$

Here $\bar{u}$ and $\bar{y}$ are the control variables, $\bar{\xi}$ and $\bar{h}$ are the state variables, and $\Omega_{\mathcal{P}_2}$ is the quadratic mapping defined in (72) associated with $\hat{\lambda}$.

Let $\bar{\chi}$ and $\bar{\chi}_h$ be the costate variables corresponding to $\bar{\xi}$ and $\bar{h}$, respectively. Note that the qualification hypothesis in Assumption 7.1 implies that $\{D\eta_j(\hat{x}_0, \hat{x}_T)\}_{j=1}^{d_\eta}$ are linearly independent. Hence any weak solution $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h})$ of (LQ) has a unique associated multiplier $\lambda^{LQ} := (\bar{\chi}, \bar{\chi}_h, \beta^{LQ})$ solution of the system that we describe next. The pre-Hamiltonian for (LQ) is

$$\begin{aligned}&\mathcal{H}[\lambda^{LQ}](\bar{\xi}, \bar{u}, \bar{y}) := \bar{\chi}(A\bar{\xi} + E\bar{u} + B_1\bar{y})\\ &+(\tfrac{1}{2}\bar{\xi}^\top Q\bar{\xi} + \bar{u}^\top E\bar{\xi} + \bar{y}^\top M\bar{\xi} + \tfrac{1}{2}\bar{u}^\top R_0\bar{u} + \bar{y}^\top J\bar{u} + \tfrac{1}{2}\bar{y}^\top R_1\bar{y}),\end{aligned} \tag{139}$$

and the endpoint Lagrangian is given by

$$\ell^{LQ}[\lambda^{LQ}](\bar{\xi}_0, \bar{\xi}_T, \bar{h}_T) := g(\bar{\xi}_0, \bar{\xi}_T, \bar{h}_T) + \sum_{j=1}^{d_\eta} \beta_j^{LQ} D\eta_j(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{h}_T). \tag{140}$$

The costate equation for $\bar{\chi}$ is

$$-\dot{\bar{\chi}}_t = D_{\bar{\xi}}\mathcal{H}[\lambda^{LQ}] = \bar{\chi}A + \bar{\xi}^\top Q + \bar{u}^\top E + \bar{y}^\top M, \tag{141}$$

with the boundary conditions

$$\begin{aligned} \bar{\chi}_0 = {} & -D_{\bar{\xi}_0}\ell^{LQ}[\lambda^{LQ}] \\ = {} & -\rho_0^{LQ}\left[\bar{\xi}_0^\top D^2_{x_0^2}\ell + (\bar{\xi}_T + B_T\bar{h})^\top D^2_{x_0x_T}\ell\right] - \textstyle\sum_{j=1}^{d_\eta}\beta_j^{LQ} D_{x_0}\eta_j, \end{aligned} \tag{142}$$

(143)

$$\begin{aligned} \bar{\chi}_T = {} & D_{\xi_T}\ell^{LQ}[\lambda^{LQ}] \\ = {} & \rho_0^{LQ}\left[\bar{\xi}_0^\top D^2_{x_0x_T}\ell + (\bar{\xi}_T + B_T\bar{h})^\top D^2_{x_T^2}\ell\right] + \bar{h}^\top C_T + \textstyle\sum_{j=1}^{d_\eta}\beta_j^{LQ} D_{x_T}\eta_j. \end{aligned}$$

For the costate variable $\bar{\chi}_h$ we get the equation and endpoint conditions

$$\dot{\bar{\chi}}_h = 0, \tag{144}$$

$$\bar{\chi}_{h,0} = 0, \tag{145}$$

$$\bar{\chi}_{h,T} = D_{\bar{h}}\ell^{LQ}[\lambda^{LQ}]. \tag{146}$$

Hence, $\bar{\chi}_h \equiv 0$ and thus (146) yields

(147)

$$0 = \rho_0^{LQ}\left[\bar{\xi}_0^\top D^2_{x_0x_T}\ell\, B_T + (\bar{\xi}_T + B_T\bar{h})^\top (D^2_{x_T^2}\ell\, B_T + C_T^\top)\right] + \sum_{j=1}^{d_\eta}\beta_j^{LQ} D_{x_T}\eta_j B_T.$$

The stationarity with respect to the control $(\bar{u}, \bar{y})$ implies

$$0 = \mathcal{H}_{\bar{u}} = \bar{\chi}D + \xi^\top E^\top + \bar{u}^\top R_0 + \bar{y}^\top J, \tag{148}$$

$$0 = \mathcal{H}_{\bar{y}} = \bar{\chi}B_1 + \bar{\xi}^\top M^\top + \bar{u}^\top J^\top + \bar{y}^\top R_1. \tag{149}$$

Notation: Denote by (LQS) the set of equations consisting of (137)-(138), (141)-(143),(147)-(149).

Note that if the uniform positivity (77) holds, then (LQ) has a unique optimal solution $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) = 0$. Besides, in view of Corollary 6.3, the strengthened Legendre-Clebsch condition holds for (LQ) at $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) = 0$. Hence, the unique local optimal solution of (LQ) is characterized by its first order optimality system (LQS). This leads to the following result.

Proposition 8.5. *If the uniform positivity in* (77) *holds, the system (LQS) has a unique solution* $(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) = 0$.

8.3. **The transformation.** Given $(\bar{x},\bar{u},\bar{v},\bar{p},\bar{\beta}) \in \mathcal{W}\times W^{1,\infty}\times \mathbb{R}^{d_\eta,*}$, define
(150)
$$\bar{y}_t := \int_0^t \bar{v}_s ds,\ \bar{\xi} := \bar{x} - B\bar{y},\ \bar{\chi} := \bar{p} + \bar{y}^\top C,\ \bar{\chi}_h := 0,\ \bar{h} := \bar{y}_T,\ \beta_j^{LQ} := \bar{\beta}_j.$$

Lemma 8.6. *The one-to-one linear mapping*

$$(\bar{x},\bar{u},\bar{v},\bar{p},\bar{\beta}) \mapsto (\bar{\xi},\bar{u},\bar{y},\bar{h},\bar{\chi},\bar{\chi}_h,\beta^{LQ}) \tag{151}$$

defined by (150) *converts each solution of (LS) into a solution of (LQS).*

Proof. It is an easy extension of Lemma 7.1 in [7]. □

We shall now go back to the convergence Theorem 8.1.

Proof. [of Theorem 8.1] Let $(\bar{x},\bar{u},\bar{v},\bar{p},\bar{\beta})$ be a solution of (LS), and let $(\bar{\xi},\bar{u},\bar{y},\bar{h},\bar{\chi},\bar{\chi}_h,\beta^{LQ})$ be defined by the transformation in (150). Hence we know by Lemma 8.6 that $(\bar{\xi},\bar{u},\bar{y},\bar{h},\bar{\chi},\bar{\chi}_h,\beta^{LQ})$ is solution of (LQS). As it has been already shown in Proposition 8.5, condition (77) implies that the unique solution of (LQS) is 0. Hence $(\bar{\xi},\bar{u},\bar{y},\bar{h},\bar{\chi},\bar{\chi}_h,\beta^{LQ}) = 0$ and thus $(\bar{x},\bar{u},\bar{v},\bar{p},\bar{\beta}) = 0$. Conclude that the unique solution of (LS) is 0. This yields the injectivity of $\mathcal{S}'$ at $\hat{\nu}$, and hence the result follows.

□

Remark 8.7 (The shooting algorithm for the control constrained case)*.* We claim that the formulation of the shooting algorithm above and the proof of its local convergence can be done also for problems where the controls are subject to bounds of the type

$$0 \le u_t \le 1, \quad 0 \le v_t \le 1, \quad \text{a.e. on } [0,1]. \tag{152}$$

This extension should follow the procedure in Section 8 of [7].

9. Conclusion

We studied optimal control problems in the Mayer form with systems that are affine in some components of the control variable. A set of 'no gap' necessary and sufficient second order optimality conditions is provided. These conditions apply to a weak minimum and do not assume the uniqueness of multipliers. For qualified solutions, we proposed a shooting algorithm and proved that its local convergence is guaranteed by the sufficient condition above-mentioned.

There are several issues in this direction of investigation that remain open. For instance, one can think of the study of other type of minimum, like Pontryagin or strong. Other possible task is the optimality of bang-singular solutions, that had not yet been deeply looked into but show to be useful in practice. Therefore, the results presented can be pursued by many interesting extensions.

References

[1] R.A. Adams. *Sobolev spaces.* Academic Press, New York, 1975.

[2] A.A. Agrachev and Y.L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences.*

[3] V.M. Alekseev, V.M. Tikhomirov, and S.V. Fomin. *Optimal Control.* Nauka, Moscow, 1979. [in Russian].

[4] M.S. Aronna. Partially affine control problems: second order conditions and a well-posed shooting algorithm. *INRIA Research Rapport Nr. 7764*, October 2011.

[5] M.S. Aronna, J. F. Bonnans, A. V. Dmitruk, and P.A. Lotito. Quadratic conditions for bang-singular extremals. *Numerical Algebra, Control and Optimization, AIMS, special issue in honor of Helmut Maurer*, 2(3):511–546, 2012.

[6] M.S. Aronna, J.F. Bonnans, and P.A. Lotito. Continuous time optimal hydrothermal scheduling. *Proceedings of the International Conference on Engineering Optimization*, 2008.

[7] M.S. Aronna, J.F. Bonnans, and Martinon P. A shooting algorithm for problems with singular arcs. *INRIA Research Rapport Nr. 7763*, October 2011.

[8] D.M. Azimov. Active sections of rocket trajectories. A survey of research. *Avtomat. i Telemekh.*, (11):14–34, 2005.

[9] D.J. Bell and D.H. Jacobson. *Singular Optimal Control Problems.* Academic Press, 1975.

[10] D.S. Bernstein and V. Zeidan. The singular linear-quadratic regulator problem and the Goh-Riccati equation. In *Proceedings of the IEEE Conference on Decision and Control*, volume 1, pages 334–339, 1990.

[11] F. Bonnans, J. Laurent-Varin, P. Martinon, and E. Trélat. Numerical study of optimal trajectories with singular arcs for an Ariane 5 launcher. *Journal of Guidance, Control, and Dynamics*, 32(1):51–55, 2009.

[12] B. Bonnard and J.-B. Caillau. Introduction to nonlinear optimal control. In *Advanced topics in control systems theory*, volume 328 of *Lecture Notes in Control and Inform. Sci.*, pages 1–60. Springer, London, 2006.

[13] H. J. Bortolossi, M. V. Pereira, and C. Tomei. Optimal hydrothermal scheduling with variable production coefficient. *Math. Methods Oper. Res.*, 55(1):11–36, 2002.

[14] H. Brézis. *Analyse fonctionnelle.* Masson, Paris, 1983.

[15] A.E. Bryson, Jr. and Y.C. Ho. *Applied optimal control.* Hemisphere Publishing Corp. Washington, D. C., 1975. Optimization, estimation, and control, Revised printing.

[16] R. Bulirsch. Die Mehrzielmethode zur numerischen Losung von nichtlinearen Randwertproblemen und Aufgaben der optimalen Steuerung. *Report der Carl-Cranz Gesellschaft*, 1971.

[17] D.I. Cho, P.L. Abad, and M. Parlar. Optimal production and maintenance decisions when a system experience age-dependent deterioration. *Optimal Control Applications and Methods*, 14(3):153–167, 1993.

[18] A.V. Dmitruk. Quadratic conditions for a weak minimum for singular regimes in optimal control problems. *Soviet Math. Doklady*, 18(2), 1977.

[19] A.V. Dmitruk. Jacobi-type conditions for the problem of Bolza with inequalities. *Math. Notes*, 35(6):427–435, 1984.

[20] A.V. Dmitruk. Jacobi type conditions for singular extremals. *Control & Cybernetics*, 37(2):285–306, 2008.

[21] A.V. Dmitruk. Quadratic order optimality conditions for extremals completely singular in part of controls. *Operations Research Proceedings 2010, Selected Papers of the Annual International Conference of the German Operations Research Society*, 0:341–346, 2011.

[22] A.V. Dmitruk and K.K. Shishov. Analysis of a quadratic functional with a partly singular Legendre condition. *Moscow University Comput. Math. and Cybernetics*, 34(1):16–25, 2010.

[23] A.Ya. Dubovitskii and A.A. Milyutin. Extremum problems with constraints. *USSR Comp. Math. and Math. Phys.*, 5(3):1–80, 1965.

[24] R. Fletcher. *Practical methods of optimization. Vol. 1.* John Wiley & Sons Ltd., Chichester, 1980. Unconstrained optimization, A Wiley-Interscience Publication.

[25] R.H. Goddard. *A Method of Reaching Extreme Altitudes*, volume 71(2) of *Smithsonian Miscellaneous Collections*. Smithsonian institution, City of Washington, 1919.

[26] B.S. Goh. Necessary conditions for singular extremals involving multiple control variables. *J. SIAM Control*, 4:716–731, 1966.

[27] B.S. Goh. *Necessary Conditions for the Singular Extremals in the Calculus of Variations*. University of Canterbury, 1966.

[28] B.S. Goh. The second variation for the singular Bolza problem. *J. SIAM Control*, 4(2):309–325, 1966.

[29] B.S. Goh. Optimal singular control for multi-input linear systems. *J. Math. Anal. Appl.*, 20:534–539, 1967.

[30] B.S. Goh. Compact forms of the generalized legendre-clebsch conditions and the computation of singular control trajectories. In *Proceedings of the American Control Conference*, volume 5, pages 3410–3413, 1995.

[31] B.S. Goh. Optimal singular rocket and aircraft trajectories. In *Control and Decision Conference, 2008. CCDC 2008*, pages 1531 –1536, 2008.

[32] M.R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.*, 1(4):525–581, 1951.

[33] D.G. Hull. Optimal guidance for quasi-planar lunar ascent. *J. Optim. Theory Appl.*, 151(2):353–372, 2011.

[34] H.J. Kelley. A second variation test for singular extremals. *AIAA Journal*, 2:1380–1382, 1964.

[35] H.J. Kelley, R.E. Kopp, and H.G. Moyer. Singular extremals. In *Topics in Optimization*, pages 63–101. Academic Press, New York, 1967.

[36] P. Kunkel and V. Mehrmann. *Differential-algebraic equations: Analysis and numerical solution*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2006.

[37] S. Kurcyusz and J. Zowe. Regularity and stability for the mathematical programming problem in Banach spaces. In *Applied Mathematics and Optimization*, pages 49–62. Springer, New York, 1979.

[38] D. F. Lawden. *Optimal trajectories for space navigation*. Butterworths, London, 1963.

[39] H. Maurer. Numerical solution of singular control problems using multiple shooting techniques. *J. of Optim. Theory Appl.*, 18(2):235–257, 1976.

[40] H. Maurer and W. Gillessen. Application of multiple shooting to the numerical solution of optimal control problems with bounded state variables. *Computing*, 15(2):105–126, 1975.

[41] H. Maurer, J.-H. Kim, and G. Vossen. On a state-constrained control problem in optimal production and maintenance. In Christophe Deissenberg, Richard Hartl, H. M. Amman, and B. Rustem, editors, *Optimal Control and Dynamic Games*, volume 7 of *Advances in Computational Management Science*, pages 289–308. Springer US, 2005.

[42] H. Maurer and N.P. Osmolovskii. Second order sufficient optimality conditions for a control problem with continuous and bang-bang control components: Riccati approach. In *System modeling and optimization*, volume 312 of *IFIP Adv. Inf. Commun. Technol.*, pages 411–429. Springer, Berlin, 2009.

[43] A.A. Milyutin. On quadratic conditions for an extremum in smooth problems with a finite-dimensional range. *Methods of the Theory of Extremal Problems in Economics*, pages 138–177, 1981.

[44] A.A. Milyutin and N.P. Osmolovskii. Calculus of variations and optimal control. *American Mathematical Society*, 1998.

[45] H.J. Oberle. Numerische Behandlung singulärer Steuerungen mit der Mehrzielmethode am Beispiel der Klimatisierung von Sonnenhäusern. *PhD thesis. Technische Universität München*, 1977.

[46] H.J. Oberle. On the numerical computation of minimum-fuel, Earth-Mars transfer. *J. Optimization Theory Appl.*, 22(3):447–453, 1977.

[47] H.J. Oberle. Numerical computation of singular control problems with application to optimal heating and cooling by solar energy. *Appl. Math. Optim.*, 5(4):297–314, 1979.

[48] H.J. Oberle. Numerical computation of singular control functions in trajectory optimization problems. *J. Guidance Control Dynam.*, 13(1):153–159, 1990.

[49] H.J. Oberle and K. Taubert. Existence and multiple solutions of the minimum-fuel orbit transfer problem. *J. Optim. Theory Appl.*, 95(2):243–262, 1997.

[50] N.P. Osmolovskii. *Theory of higher order conditions in optimal control*. PhD thesis, Moscow Civil Engineering Institute, 1988.

[51] H.J. Pesch. A practical guide to the solution of real-life optimal control problems. *Control Cybernet.*, 23(1-2):7–60, 1994. Parametric optimization.

[52] L. Pontryagin, V. Boltyanski, R. Gamkrelidze, and E. Michtchenko. *The Mathematical Theory of Optimal Processes*. Wiley Interscience, New York, 1962.

[53] H.M. Robbins. A generalized Legendre-Clebsch condition for the singular case of optimal control. *IBM J. of Research and Development*, 11:361–372, 1967.

[54] E. Trélat. Optimal Control and Applications to Aerospace: Some Results and Challenges. *J. Optim. Theory Appl.*, 154(3):713–758, 2012.

M.S. Aronna, ITN Marie Curie Network SADCO at Università degli Studi di Padova, Italy

E-mail address: aronna@math.unipd.it